Firing activities of auditory cortical neurons during categorical task performance in behaving monkeys

Dissertation

zur Erlangung des akademisches Grades

doktor rerum naturallium
(Dr. rer.nat.)

genehmigt durch die Fakultät für Naturwissenschaften
der Otto-von-Guericke-Universität Magdeburg

von Elena Selezneva
geb. am 05.07.1977 in Moskau

Gutachter:	Prof. Dr. Thomas Münte
	Prof. Dr. Henning Scheich
	Prof. Dr. Yuri Alexandrov

eingereicht am: 30.05.06
verteidigt am: 11.01.07

Acknowledgment

This work was performed during the time from October 2000 till May 2006 in Leibniz Institute for Neurobiology, Department Auditory Learning and Speech, under the direction of Prof. Dr. Henning Scheich. Here, I would like to thank the people who have helped during my studies for making my experience in the lab both educational and pleasurable.

First of all I would like to thank Prof. Dr. Henning Scheich, for giving me the opportunity to do my PhD in his group, for his ideas and for his support. I am also especially thankful to my supervisor PD Dr. Michael Brosch for his guidance, encouragement and endless patience. I sincerely appreciate his suggestions and his help.

I thank Cornelia Bucks who not only provided an excellent technical assistance by behavioral and electrophysiological experiments but also taught me German and was always a good friend for me.

I would also like to thank Elena Oshurkova for moral and technical support as well as all other colleagues in Leibniz Institute for Neurobiology for the good atmosphere and their friendly cooperation.

I am greatly indebted to Prof Dr. Yuriy I. Alexandrov, Dr. Alexandr G. Gorkin and Dr. Vladimir V. Gavrilov, who gave me theoretical and practical knowledge and guided my way in science.

Table of Contents:

1. General introduction

Categorization is the act of assigning objects or events to classes (i.e., categories). By categorical perception the continuous and variable stimulation that reaches the sense organs is sorted out by the mind into discrete, distinct classes whose members come to resemble one another more than they resemble members of other categories (Harnad, 1987). It is performed countless times every day, and is among the most important and basis of all decisions. The best-known examples are color categories, relative musical pitches (rising versus falling) and stop-consonants in speech sounds. Categorization can be considered as a means of structuring the surroundings and parsing it into units that can be processed, manipulated, and stored more efficiently than continuous variation. From an ethologist view categorization has been viewed as a process of searching for the set of releasers or key stimuli that trigger a specific behavior. General behavior is based on categorization, while selecting a specific action towards a particular object is based on recognition. During perceptual categorization, unrelated components of the environments are linked up and are given salience for decision-making so that a group of individually different stimuli leads to the same response.

Categorization is not a single mental ability, but instead depends on several different abilities that use different brain structures and processes. Human studies have identified at least three different kinds of category-learning tasks (Ashby and Shawn, 2001), depending on how the categories are constructed. The neural circuitries that mediate each type of category learning are also at least partly different which was confirmed by neuropsychological studies with different patient groups and also by recent neuroimaging data.

In rule-based tasks, subjects learn the category structures via some explicit reasoning process. In this case, the optimal rule to determine the category membership is often easy to describe verbally (Ashby et al., 1998). Certainly most of standard neuropsychological categorization tasks are of this type. According to neuroimaging data (Rao et al., 1997; Elliott et al., 1999), the important structures for rule-based category learning are prefrontal cortex and basal ganglia. This data corresponds to the neuropsychological studies of category learning (Brown and Marsden, 1988; Cools et al., 1984; Robinson et al., 1980), in which was shown that individuals with frontal lobe or basal ganglia dysfunctions are impaired in rule-based tasks.

Information-integration tasks are those in which accuracy is maximized only if formation from two or more stimulus components must be integrated at some pre-decisional stage (Ashby and Gott, 1988). A neuroimaging study (Seger and Cincotta, 2002) shows striatal and lateral occipital activation in a group of subjects performed this task. According to neuropsychological data (Maddox and Filoteo, 2001; Filoteo et al., 2001), patients with striatal dysfunction are impaired by performing of information-integration tasks. Note that when a category contains only a few highly distinct exemplars, memorization is necessary. In this case, patients with temporal lobe amnesia showed also late training deficit (Knowlton et al., 1994).

In prototype distortion tasks, each category is created by first defining a category prototype and then creating the category members by randomly distorting these prototypes (Posner and Keele, 1968; Homa et al., 1981). Neither individuals with frontal lobe lesions nor individuals with disease of the basal ganglia or medial temporal lobes disease were impaired in this type of task (Knowlton et al., 1992, 1996; Kolodny, 1994; Meulemans et al., 1998). The fMRI studies show learning-related changes in the visual cortex (Reber PJ et al., 1998; Aizenstein et al., 2000). This suggests the hypothesis that learning in prototype distortion tasks depends on the perceptual representation memory system, through a perceptual learning process.

Single-cells human and animal studies have also identified several brain structures that are critical for categorical perception. Note that in this case the studies were mostly focused not on the learning of new categories, but on the categorization behavior of highly experienced subjects. Thus the category-specific activity was found in prefrontal cortex (Freedman et al., 2002, 2003; Fukushi and Sawagushi, 2005), basal ganglia (Merchant et al., 1997; Romo et al., 1995), medial temporal lobes (Kreiman et al., 2000; Hampson et al., 2004), primary (Salinas and Romo, 1998) and supplementary motor cortex (Romo et al., 1993, 1997; Isomura et al., 2003).

Many studies also addressed the question if characteristics of a category and the rules for distinguishing it from similar but different categories are learned and stored in the sensory cortex. In 1977 in inferotemporal cortex of monkeys were found cells which proved to be responsive for complex visual objects (Rolls et al., 1977). More recent study has suggested that about 25% of cells in inferotemporal cortex show some degree of category-selectivity (Vogels, 1999).

In 1996 Yoshioka and colleagues reported that the categorical color perception occurs also on early stages of visual cortex. In another single-cell study, targeting the inferotemporal cortex (Sigala and Logothetis, 2002), was shown that after category learning about 70% of neurons were selective for the category-relevant stimulus dimensions, but not for the other dimensions which did not affect category membership. In 2002 Lee and colleagues reported also that behavioral training in a target detection task changed neuronal selectivity even in the primary visual cortex.

The role of sensory cortex in categorical representation was also suggested by studies undertaken in auditory cortex. Correlation between neuronal firing and categorical boundaries was found in primary auditory cortex of anesthetized cats (Eggermont, 1995) and of awake gerbils (Wetzel et al., 1998a; Ohl et al., 2001) and macaque monkeys (Steinschneider et al., 1994, 1995).

The goal of the present study was to examine whether the firing of auditory cortex neurons reflected actually the category membership of tone steps (rising versus falling) and not merely the physical characteristics of the single tones. The study was divided into two parts. First the monkeys were trained to categorize up and down pitch direction in variable sequences of pure tones. A positive-reinforcement behavioral procedure was used and only the responses to falling frequency contours were reinforced. After the monkeys had learned this task, the recording of the neuronal activity from the auditory cortex was performed simultaneously with the task performance. Then the neuronal responses to falling frequency contour and the neuronal responses to rising frequency contour were analyzed with sets of tone sequences such that for the same neuron responses to identical tones could be compared in the two cases.

2. Behavioral study

2.1. Introduction

Relative pitch perception is one of the best-known examples of categorical perception. When a melody is transposed (i.e., absolute frequencies are changed but frequency relations are preserved), humans perceive the transposed melody as similar to the original one because the contour is identical. Perception of such relationships between frequencies is prominent for humans from early stages of development (Chang and Trehub, 1977; Trehub et al., 1984). Humans readily identify and memorize melodies by the sequential up-and-down patterning of the pitches of adjacent tones in a tune (Dowling, 1978). In the present study it was tested whether monkeys are also able to discriminate pitch relationships.

A number of studies have demonstrated that animals can use simple relational concepts like identity and oddity in auditory discrimination tasks. This has been shown in tests with acoustic signals like pure tones or frequency sweeps, in which animals had to signal whether consecutive sounds were alike or different. Such discriminations could be performed for different qualities of sounds (D'Amato and Colombo, 1985; Kojima, 1985; Wright et al., 1990; Sinnott and Kreiter, 1991; Fitch et al., 1993; Sakurai, 1994; Wetzel et al., 1998b). Simple relational concepts, however, are not sufficient for the identification of spectro-temporal sound patterns that are characterized by the ordinal relation between individual segments of the pattern rather than by specific physical properties of the individual segments. There is some evidence that non-human mammals have the perceptual capability to attend to relationships between acoustic items. For example, Wright and colleagues (2000) found that monkeys can rate well-known melodies as similar when they are transposed by an octave. Similarly, a study of Hauser and colleagues (2001) suggests that monkeys extract at least parts of the sequential structure of syllables in streams of artificial speech signals. The reason why there is still so little evidence that animals can identify spectro-temporal sound patterns based on the relationship between tones seems to be that the animals' discriminative performance of sound patterns is largely controlled by absolute physical properties of individual tones in a sequence and little, if at all, by the relation between different elements of sound patterns, as D'Amato (1988) concluded after an extensive research on monkeys and rats. Izumi (2001) showed that monkeys could discriminate sequences by the relative pitch, if they were restricted to using absolute cues. However, they were able to transfer relative pitch perception to novel sequences only within the absolute frequency range which was used by training, but this percept did not transfer to sequences outside the trained range. A similar tendency to prefer for attending the absolute rather than the relative pitch has also been observed in songbirds, which only in specific conditions, namely when absolute frequency cues were mitigated, could discriminate tone sequences based on pitch relations (Page et al., 1989; Braaten et al., 1990, Braaten and Hulse, 1993; Cynx, 1995; MacDougall-Shackleton and Hulse, 1996).

Thus the first part of the present study is addressed to the question if non-human primates can obtain ordinal relations between individual tones in a sequence and discriminate the direction of the pitch change in the wide frequency range independently of absolute physical properties of individual tones in a sequence.

2.2. Methods

2.2.1. Subjects

Two adult male cynomolgus monkeys (*Macaca fascicularis)* were used in this study. Throughout the experiments, the two monkeys were housed together in a cage, in which they had free access to dry food like pellets, bread, corn flakes, and nuts. They earned a large proportion of their water ration during the daily positive-reinforcement training sessions and received the remainder in the form of fresh fruit during and after each session and in the weekends. The daily rations were sufficient to maintain the animals at 85-95 % of their free-feeding body weights. Experiments were approved by the local committee for animal care and ethics and conformed with the rules for animal experimentation of the European Communities Council Directive (86/609/EEC).

2.2.2. Apparatus

Experiments were initially carried out in an anechoic single-walled room and were later continued in a sound-shielded double-walled room (IAC, 1202-A). The monkey was seated in a custom-made restraining chair. The front panel of the chair accommodated a red light-emitting diode, a water spout, and a touch bar. The behavioral procedure was controlled, monitored, and recorded by a computer and a video camera. Response latencies were measured with a temporal resolution of 1 µs. Acoustic signals were generated digitally with the aid of the computer, which was interfaced with an array processor (Tucker-Davis Technologies, AP2-card), at a sampling rate of 100 kHz and D/A converted to an analog signal (Tucker-Davis Technologies, DA1). The analog signal was amplified (Pioneer, A204) and coupled to a free-field loudspeaker (Jamo, D265), which was located ~1.5 m in front of the animal. The sound pressure level (SPL) was measured with a free field 1/2 inch microphone (G.R.A.S., 40AC) located close to the monkey's head and a spectrum analyzer (Rion, SA 77). The output of the sound delivering system varied ±10 dB in the frequency range of 0.2-35 kHz. At sound pressure levels used in the present study (~ 60 dB SPL), harmonic distortion was > 36 dB below the signal level.

2.2.3. Procedure

Since previous studies had found it extremely difficult to train animals to extract ordinal relations (D'Amato, 1988; Wright, 1991), the training was divided into two phases with increasing task difficulty. Stimulus properties as well as reward contingencies were adjusted carefully and gradually during the course of the training to keep the monkeys at reasonable reward rates and, thus, in a motivated and non-frustrated state.

In phase I, the monkeys were trained to respond when there was a change of the frequency in a sequence of pure tones. In phase II, the monkeys were trained to distinguish categorically an upward pitch direction from a downward pitch direction.

The general layout of the experiment is shown in Fig. 1. A trial started by turning on the light-emitting diode (LED), which was the signal for the monkey to make contact with the touch bar. After a variable period of 0.6-2 sec, such contact triggered a sequence of pure tones of two or, in phase II, three different frequencies. The monkey's task was to release the touch bar upon occurrence of the first tone of lower frequency. When they did so within a specified response interval, commencing 0-0.3 sec and ending 1.2-2 sec after onset of the stimuli, a water reward was delivered. Releasing contact at any other time prompted an immediate termination of the stimulation and a 7-sec time-out from the experiment as a mild form of punishment. The cue-light was extinguished at the end of a trial, and there was a 5-sec intertrial interval before the next trial was started. For the procedure, monkey F used his left hand and monkey B his right hand.

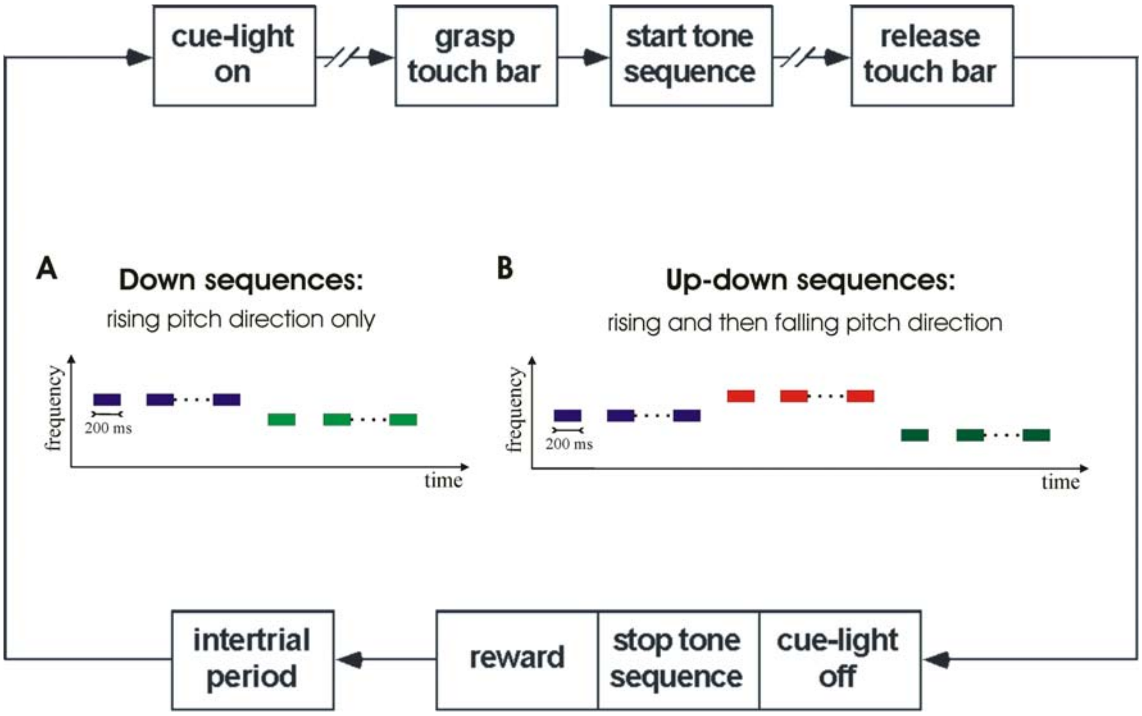

Figure 1: Visual representation of the behavioral paradigm used in the two training phases. The beginning of a trial was indicated by switching on a cue-light. If the monkey made contact with a touch bar within a specified time interval, a sequence of pure tones of different frequencies was played with a constant delay from the loudspeaker. Upon occurrence of the falling frequency contour monkeys had to release the touch bar. Responses were considered correct when the monkey responded during the response period, which commenced 0-0.3 sec and ended 1.2-2 sec after the onset of the go stimuli. **(A)** Sequences used in phase I and in parts of phase II. The first tones in the sequence had the same frequency. They were followed by tones of lower frequency. The frequency of initial tones varied randomly from trial to trial while the frequency interval (ratio between the frequencies) was constant. This type of sequences was termed “down sequences”. **(B)** Sequences used in training phase II. The first tones in the sequence were followed by a variable number of tones of higher frequency and then by tones of lower frequency. This type of sequences was termed “up-down sequences”.

The percentage of correct responses was calculated to assess the animals' performance in a session. It was defined as the number of trials with responses made within the response interval, divided by the total number of trials in which the monkey made contact with the touch bar after the cue-light had been switched on. Error trials, thus, included trials with responses before and during the presentation of the initial tones of the same frequency, trials with responses during the presentation of the tones of higher frequency (in phase II), as well as trials in which the monkey maintained contact with the touch bar after the cessation of falling frequency contour. It was considered that a monkey had learned a specific task when he made significantly more correct than incorrect responses.

The specific stimulus parameters used in the different training phases as well as the behavioral performance of the monkeys will be described in detail below. Individual training sessions lasted 2-4 hours (including pauses) during which the monkeys made 300-800 trials.

Note that prior to the present experiments, both animals were acquainted with the behavioral procedure, i.e., they learned to make contact with the touch bar for some time after illumination of the LED light, to notice sounds and to release the touch bar upon occurrence of a go stimulus (noise or clicks). In this pre-training phase monkeys learned also to discriminate acoustic items of different sound quality (noise bursts versus clicks trains). Both monkeys scored ~90% of correct performance in these tasks.

2.3. Training Phase I: Detection of a pitch change in a sequence of pure tones

2.3.1. Methods

In phase I, tone sequences consisted of pure tones of two different frequencies. The initial tones were all set at the same frequency. They were followed by tones which were all set at another frequency, always below that of the initial tones. This type of sequences was termed “down sequences”. The initial frequency varied randomly from trial to trial while the frequency interval (ratio between frequencies) was constant. The number of initial tones varied during the training phase and also during one training session. Thus, monkeys listened to a sequence of repeating tones, in which there was one direction of a pitch change, namely downwards, and were required to respond to this change. All tones in a sequence had the same intensity, which ranged between 55 and 65 dB in different sessions. Tone duration and intertone intervals were 300 ms.

2.3.2. Results

In the first few training sessions of phase I, a small number of initial tones was used. The frequencies for both the initial tones and the tones of lower frequency were fixed. The frequencies were 2000 Hz and 670 Hz in monkey F and 800 Hz and 262 Hz in monkey B. Initially both monkeys responded at chance to these sequences, indicated by a drop of the percentage of correct responses and the variation of reaction times.

To force animals to cope with this task, the variability of the number of initial tones were increased to maximally 10. After being trained with the sequences of pure tones for 8 sessions monkey F resumed to respond to the occurrence of the stimuli of lower frequency. Monkey B required 7 sessions to be able to respond to the falling frequency contour in sequences with pure tones. The percentage of correct responses was ~70 % for both monkeys. The mean reaction time was 600 ± 80 ms by monkey F and 550 ± 150 ms by monkey B.

At this time it was not clear whether monkeys actually detected the changing of frequency to perform the task or whether they simply responded to the absolute frequency of the tones. Therefore the frequency of the tones was next varied from trial to trial. The frequency ratio between tones maintained constant.

In monkey F, the trial-to-trial variability of the initial frequency could be gradually increased up to 0.7-8.5 kHz during the following 22 sessions without observing a deterioration of his performance. Despite this considerable frequency variability used at this training stage it was discovered that the monkey attended largely to the absolute frequency of the falling frequency contour. By using sequences with relatively low initial frequencies (0.7-2.4 kHz), the monkey failed to faithfully respond to the occurrence of the stimuli of lower frequency. The same behavior was observed in the other monkey. Because of this tendency both monkeys were trained, first session-wise and then block-wise, during the following training sessions with tones within a low or a high frequency range only.

In parallel the frequency ratio between tones varied between 0.8 and 1.6 octaves. After another 23 session in monkey F and 30 sessions in monkey B, a wide frequency range between 0.5-20 kHz was tried again. Monkeys now responded to the occurrence of the frequency change and did so independent of the frequency of the tones in the sequence and of the number of initial tones. In monkey F, the percentage of correct responses was at 84 % and the reaction time was 680 (± 200) ms. Monkey B responded correctly in 80 % of the trials and had a reaction time of 560 (± 150) ms.

These results indicate that both monkeys had learned to detect a pitch change in a sequence of pure tones. The stimulus material ensured that the monkeys could not employ other cues for a successful detection of the falling frequency contour, like the absolute frequency or the number of initial tones.

2.4. Training phase II: Discrimination of pitch direction

2.4.1. Methods

In phase II, to the sequences used in phase I were added a series of tones of higher frequency, hence forming sequences with two pitch changes. As these tones were added after initial tones, the first frequency change was in upward direction and the second change was in downward direction. This type of sequences was termed up-down. Thus two types of tone sequences were used: down sequences consisting of two series of repeating tones and up-down sequences consisting of three series of repeating tones of different frequency. Monkeys had to desist from responding to the rising frequency contour (when the frequency of the tones changed upwards) and were required to respond only to the falling frequency contour (when the frequency of the tones changed downwards).

2.4.2. Results

The training of discriminating pitch direction turned out to be quite difficult. Different procedures were tried during the course of 199 sessions in monkey F and 211 sessions in monkey B. Both monkeys were trained in parallel over wide periods and, with a few exceptions, most procedures were performed on both monkeys. Procedures that were apparently unsuccessful will not be described.

In monkey F, phase II started by adding 2-3 tones whose frequency was just slightly above the frequency of initial tones (0.01-0.02 octaves). The number of initial tones was reduced to 2-7 such that the total number of stimuli in a sequence was the same as at the end of phase I. All other parameters were as at the end of phase I: the frequency of initial tones varied between 0.5 and 20 kHz and the second frequency change was 1 octave. With this small frequency interval by the first frequency change, the performance of monkey F was indistinguishable from that at the end of phase I, in which there was only one pitch change. This suggests that the monkey treated the sequences with one or two frequency changes as alike. The same behavior was observed in monkey B who listened to similar sequences. However, after increase the first frequency interval both monkeys tended to respond to the first frequency change and did not wait until the second change. Consequently the percentage of correct responses of the monkeys declined to levels well below ~ 70 % and they nearly refused to participate constructively in the experiment.

To aid the monkeys to not respond to the first frequency change, the intensity of the tones of lower frequency was increased by 5-10 dB such that these tones were louder than the initial tones and tones of higher frequency. All other sequence parameters, including the interval of first frequency change, were not changed. Within 9 sessions with this additional cue, monkey F learned to desist from responding to the first frequency change and to respond to the second frequency change. Therefore the intensity of the tones of lower frequency could be decreased to the intensity of the other tones. Monkey F now responded correctly to 75 % of the sequences. Monkey B underwent a similar training schedule with slightly different sequence parameters and exhibited similar problems. Nonetheless he also learned this task within 14 sessions and ultimately scored 78 % correct.

In the next sessions (84 for monkey F and 96 sessions for monkey B), the first frequency interval was slowly incremented until it was similar to or larger than the second interval. The progress in incrementing the first frequency interval became faster when the variability of the number of higher frequency tones was increased. To keep the average sequence duration roughly constant, the variability of the number of initial tones was decreased in parallel. It was tried also to use sequences with tones of shorter duration and shorter intertone intervals (down to 50 ms each) while increasing their number such that the time range within which the two frequency changes occurred remained constant. This modification had no effect on the monkeys' performance so finally the tone durations and the intertone intervals were 200 ms for monkey F and 300 ms for monkey B.

At this training stage both monkeys could discriminate the sequences either by attending to the pitch direction or by attending to the ordinal position of the pitch change, i.e., by refraining to respond to the first change and responding to the second change. Therefore down and up-down sequences, i.e., sequences with one ore two frequency changes were next alternated, in blocks of 10-40 trials. As a result of these modifications both monkeys responded only occasionally and thus at chance to sequences with one frequency change during the following sessions. This indicates that they responded to the ordinal position of the pitch change. Nevertheless, within 19 sessions monkey F gradually learned to respond equally often to a falling frequency contour, whether or not it was preceded by a rising frequency contour. Monkey B took 21 sessions to learn this task.

Next down and up-down sequences were alternated randomly, instead of block-wise. This modification again resulted in a deterioration of the performance of both monkeys, suggesting that the animals did not discriminate the sequences according to the pitch direction but rather attended to the ordinal position of the frequency change and had learned to respond to the first or to the second change according to the temporary stimulus statistics: they preferentially responded to the first frequency change in blocks of trials with one frequency change and to the second frequency change in blocks of trials with two frequency changes. Fortunately, this behavior changed during the following sessions, in which the monkeys were trained with sequences in which the number of tones of higher frequency varied randomly between 0, 3, 4, or 5. After having worked with these sequences for 29 sessions monkey F's performance was largely independent of the number of higher frequency tones. Monkey B took 21 sessions to reach a similar performance.

Although both monkey were successful in task performance, it turned out that they discriminated the sequences by a combination of two other cues. At this training stage the sequence parameters used for monkey F were the following: the frequency of initial tones varied between 0.7 and 8 kHz. The frequency of the tones succeeding the initial tones was 1.1 octaves lower in down sequences and 1.6 octaves higher in up-down sequences. Consequently, the absolute frequency of the tones succeeding the initial tones was highly related to pitch direction: the second frequency was in the low range (327-3732 Hz) in sequences with a downward pitch direction only. It was in a high frequency range (2122-24252 Hz) when there were two pitch changes. The situation was similar in monkey B for whom were presented an initial tone frequency range of 0.7-8 kHz and a rising frequency interval of 0.7 octaves and a falling frequency interval of -1.5 octaves. Because of the selection of these sequence parameters the monkeys could simply attend to the frequency of the tones succeeding the initial tones and respond when this frequency was low and maintain contact when this frequency was high and wait with the response until the occurrence of the second frequency change. That the monkeys actually applied this strategy was indicated by finding that they confused sequences in which the frequency of the tones succeeding the initial tones was ambiguous in predicting whether the first or the second frequency change was the rewarded one. The use of this strategy again reflects the preference of monkeys to attend to the absolute frequency of the tones.

Therefore the stimulus material was adjusted once again such that the frequency of the first series of tones in the sequence, as well as that of the second series of tones, gave less hints as to whether the first or the second frequency change was rewarded. This adjustment was done by using different frequency ranges for initial tones in down and up-down sequences: in down sequences low frequencies of initial tones were omitted, while in up-down sequences this was done for high frequencies. The shift between the two frequency ranges of initial tones was equal to the shift between the two frequency ranges used for the second series of tones in the sequence. Consequently the frequency of the tones succeeding the initial tones was much more ambiguous in predicting which frequency change was rewarded. The correlation between absolute frequency and pitch direction was further diminished by increasing the total frequency range of initial tones to 0.6-13.2 kHz and by decreasing the size of the intervals between the tones. The rising frequency contour was 0.6 and 0.8 octaves in monkey F and B, respectively, whereas the falling frequency contour was -0.9 and -1.1 octaves.

With these modified sequences, it took monkey F 56 sessions and monkey B 57 sessions to change their decision criterion. To find out when the monkeys actually discriminated pitch direction, following statistical testing was performed. First the frequency range of initial tones in down and up-down sequences was divided into 8 bins each such that all bins had the same sample size. Then it was calculated how many trials were required such that a statistical test revealed with a power > 95 % that the percentage of correct responses was > 15 % above the percentage of incorrect responses for each of the 8 frequency bins and for both types of sequences. This calculation yielded that always data from 6-9 consecutive sessions had to be combined to reach this testing power. Last for each of these chains of sessions was determined whether the monkey responded, for all 8 frequency bins and for both types of sequences, significantly ($p < 0.05$) more often during the response period than outside this period. The response period started 240 ms after onset of the tones of lower frequency and had a duration of 1200 ms.

This analysis showed that monkey F discriminated sequences with a downward pitch direction from sequences with an upward pitch direction in the last 9 of the 56 sessions of the final training stage. He did so independent of the frequency of the tones and independent of the position of the downward pitch direction in the sequence (Fig. 2). In these 9 sessions, monkey F scored an average of 72 % correct and had a reaction time of 805 (± 236) ms. Monkey B also succeeded in the last 7 of the 57 sessions to discriminate pitch direction. He scored at 73 % correct and had a reaction time of 741 (± 184) ms.

The error rate of ~ 25 % suggests that some of the monkeys' responses were triggered by cues other than pitch direction. One likely cue was still the absolute frequency, indicated by observing that the percentage of correct responses rate of both monkeys varied with the frequency of the tones (Fig. 2). This shows

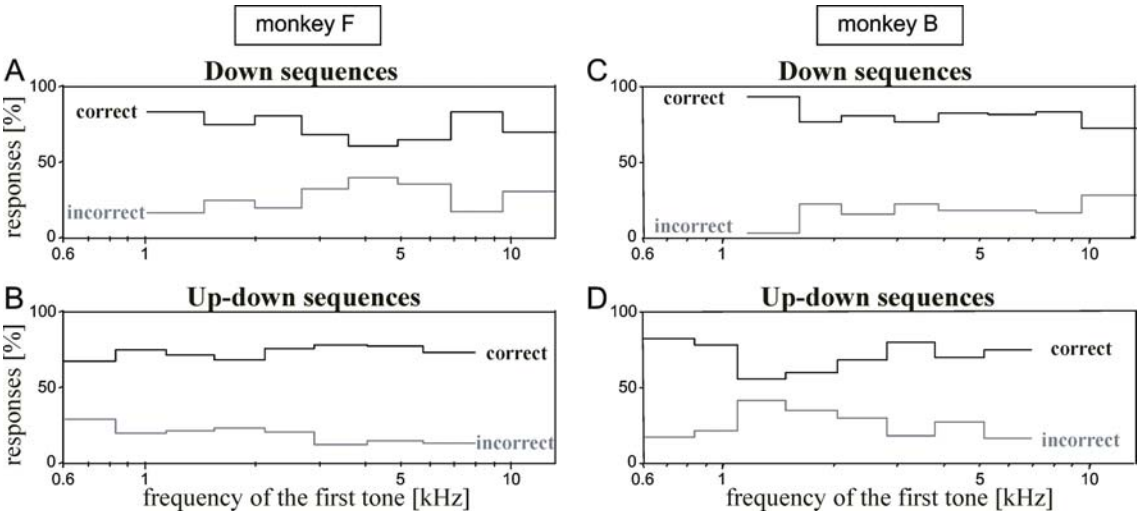

Figure 2: Monkeys discriminate pitch direction categorically. Left column shows data from monkey F, right column from monkey B. **(A, C)** Dependence of responses on the frequency of the first tone, computed for down sequences. The frequency range of initial tones was divided into 8 bins such that each bin contained the same number of cases (N = 1350 in monkey F and N = 1232 in monkey B). The percentage of correct responses was defined as the number of responses that occurred within the response period, commencing 240 ms after onset of the tone of lower frequency and lasting 1200 ms, divided by the total number of trials. Responses occurring outside the response period were considered incorrect. **(B, D)** Dependence of responses on the frequency of the first tone, computed for up-down sequences. The number of higher frequency tones varied between 2 and 6 in monkey F and 3 and 6 in monkey B. Total number of trials was 2964 in monkey F and 3015 in monkey B. Data for monkey F were from the last 9 sessions of training phase II. Data for monkey B were from the last 7 sessions of training phase II.

again the monkeys' preference to attend to physical properties of individual acoustic items. Another cue used by the monkeys was counting the number of acoustic items or waiting for a specific period of time after commencement of the tone sequence - abilities previously described in rats (Davis and Albert, 1986). In the present study, this is indicated by observing that the reward rate of the two monkeys varied with the number of higher frequency tones (Fig. 3). Both monkeys scored best when the sequences contained 3-4 tones of higher frequency. The counting or timing cue is also suggested by the type of errors the monkeys made. In down sequences, most of their errors were maintaining contact during the entire presentation of the sequences. By contrast, in up-down sequences, most errors were premature responses and the error rate increased with the number of higher frequency tones.

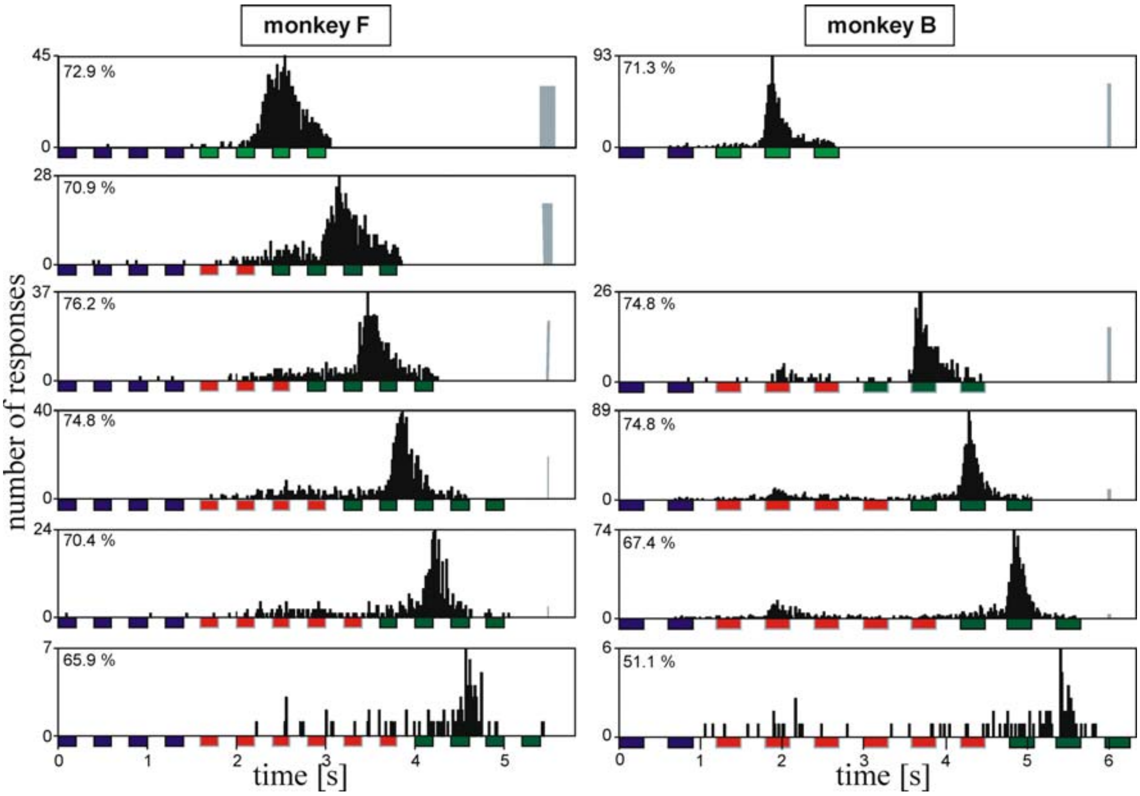

Figure 3: Distribution of response times in the pitch direction task for different number of tones of higher frequency. The tone sequences started at 0 ms. Blue rectangles underneath the x-axis symbolize initial tones, red rectangles symbolize higher frequency tones, and green rectangles symbolize lower frequency tones. Duration of tones and silent intervals were 200 ms in monkey F (left column) and 300 ms in monkey B (right column). Bin size of response times was 20 ms. Responses that occurred after the response period (240 - 1200 ms after onset of the go tones) are indicated by gray bars. Data are from the same training sessions as used for Fig. 2. Numbers give the percentage of correct responses.

2.5. Discussion

Using sequences of tones with variable frequency, the present study demonstrates that non-human primates can discriminate ascending from descending pitch contours, independent of the absolute frequency of the tones and of the ordinal and temporal position of the pitch change in the sequence. This indicates that monkeys can develop a behavioral strategy during training in which they base their discriminations on the relationships between the tones in the sequence.

The relational concept required for a successful performance in the current experiments is qualitatively different from a simple relational concept like oddity or identity. When attending to oddity subjects merely have to realize that some characteristics of the tones in the sequence have changed. With this strategy the subjects of the present study would always have responded to the first frequency change, independent of its pitch direction and, thus, subjects hardly ever would have heard the second frequency change. In order to receive a reward, subjects had to extract the ordinal relationship between consecutive stimuli, that is, they had to determine whether the succeeding tone had a higher or a lower pitch. This task is equivalent to recognizing non-auditory but categorical relationships like 'greater than/ equal to/ smaller than', 'before/after', or 'in front of/behind'.

The present data demonstrate that the concept of pitch contours exists already in non-human primates. This suggests that the idea of musical universals may not be specific only to humans. Nonetheless, there seem to be substantial interspecies differences with regard to their perceptual preference. In humans, tunes are identified and memorized mostly according to their pitch contour, i.e., to the sequential up-and-down patterning between adjacent notes, whereas the absolute pitch of the tones of a tune are not readily memorized (Dowling, 1978). Animals, in contrast, seem to attend primarily to features of the individual elements of sequential stimuli. This is exemplified by the current experiments, in which monkeys in different training stages tended to respond to the absolute pitch rather than to tone relations. Similar results have been reported in other animal studies on serial discriminations of acoustic signals (D'Amato, 1988; Page et al., 1989; Braaten et al., 1990; Braaten and Hulse, 1993; Izumi, 2001). The preference to attend to features of individual elements of tone sequences is also indicated by the difficulties exhibited by the two monkeys in phase I and the total number of sessions required for learning this task. Although both monkeys were previously trained to detect the change between noise bursts and click trains, the same/different rule, acquired in this task, was not generalized to other stimulus material. In previous studies, generalization of a non-matching concept for acoustic stimuli beyond the training context has been found to be very difficult for monkeys (D'Amato and Salmon, 1984; D'Amato and Colombo, 1985; Kojima, 1985; Wright et al., 1990; Wright, 1991) and experiments successful only when the training procedure and the stimulus material were carefully controlled.

That monkeys can make use of the concept of pitch contours in an experimental setting seems to depend on the training procedure. Although the presented training procedure was eventually successful it is not necessarily the fastest and the only way of training animals to discriminate pitch direction. A more careful selection of sequence parameter may help avoiding that animals use cues other than those desired by the experimenter. This was the case in phase II, when pitch direction was correlated with the absolute frequency of the tones in the sequences and, thus, monkeys could base their decision on the frequency cue (which they prefer) rather than on the relational cue. Another possibility to accelerate the training may be to use behavioral procedures different from go/nogo paradigms or different reward schedules. Indeed, a recent study, using a habituation-dishabituation paradigm, showed that cotton-top monkeys could discriminate the serial order of syllables in artificial languages (Hauser et al., 2001). However, the training to discriminate pitch contours may take longer or may even be unsuccessful if monkeys are confronted with complex sequences immediately after initial behavioral shaping. This is suggested by a third monkey who was also trained to recognize frequency changes (corresponding to phase I of the current study). After initial behavioral shaping this monkey was immediately exposed to tone sequences in which the frequency of initial tones varied over a wide range from trial to trial. The monkey did not show an indication of learning to respond to the frequency change during a period of 106 sessions (52200 trials), until a training procedure was started equivalent to the procedure performed with the subjects of the current study.

Thus the present experiment shows that monkeys can establish the categorical concept of pitch contours in a laboratory setting. This suggests that monkeys can be used as a model for humans to study the representation of pitch contours in auditory cortex at the single cell level.

3. Electrophysiological study

3.1. Introduction

A widely held assumption is that auditory cortex, like other early sensory cortical areas, is unimodal and primarily involved in the processing of sounds and that the auditory modality is eventually integrated with other modalities in specific brain structures (Stein and Meredith, 1993). From animal studies, however, it has long been known that neurons in the auditory cortex are not only 'feature detectors' but their activity dependents also of behavioral context and meaning of stimuli. It was shown for example that the responses evoked by the auditory stimuli in the performance condition are generally greater then in the nonperformance or anaesthetized condition (Beaton and Miller, 1975; Pfingst et al., 1977; Miller at al., 1980). The increasing of evoked response to the conditioned stimulus after training was reported for neurons in the secondary and primary auditory areas of cats (Diamond and Weinberger, 1986, 1989; Witte and Kipke, 2005), monkeys (Gottlieb et al., 1989; Ahissar et al., 1992, 1998; Blake et al., 2002), ferrets (Fritz et al., 2003, 2005) etc. Other studies reported that responses are modulated by auditory short term memory (Gottlieb et al., 1989; Sakurai, 1994), long term memory (Recanzone et al., 1993; Suga and Ma., 2003; Beitel et al., 2003; Weinberger, 2004), stimulus anticipation (Hocherman et al., 1981), attention (Hubel et al., 1959), audio-motor association (Vaadia et al., 1982; Durif et al., 2003), eye position (Werner-Reiss et al., 2003) and vocal production (Müller-Preuss and Ploog, 1981; Eliades and Wang, 2003, 2005). Aside from non-auditory modulations of auditory responses it has been found that neurons in auditory cortex can respond to somatosensory (Schroeder et al., 2001; Fu et al., 2003) or to visual stimuli alone in normal (Schroeder and Foxe, 2002) or in experimentally cross-wired animals (Sur et al., 1990). Indication for cross-modal activation of auditory cortex also comes from non-invasive imaging studies in professional musicians (Bangert et al., 2001) and in deaf subjects (Finney et al., 2001).

The involvement of the auditory cortex in categorization processes was also demonstrated by a number of human and animal studies. In 2001 Näätänen described the 'primitive intelligence' of auditory cortex that includes besides other functions also the function of permanent categorical sound perception (e.g. key- and instrument-independent melody recognition and speech perception). Some evidences supporting this cognitive role of the auditory cortex arise from investigations of human EEG using the mismatch negativity (MMN). For example, Pavalainen and colleagues (1999) presenting sequences of pairs of ascending tones, of which some were infrequently replaced by a pair of descending tones, found that the descending pairs elicited MMN in the event-related potential, which occurred 100-200 ms after the onset of the second tone in the descending pair. They suggested that the brain mechanisms underlying MMN can genuinely extract the direction of a frequency step, independently of both the absolute frequency of the pair and the within-pair frequency ratio. In other study, Phillips and colleagues (2000) using an adapted oddball paradigm showed that representation of discrete phonological categories occurs in the human auditory cortex. They demonstrated the all-or-nothing property of phonological category membership and enhanced acoustic discrimination at or near phonetic category boundaries. Recent fMRI studies also confirmed the role of auditory cortex in categorical perception (Zaehle et al., 2004; Pekkola et al., 2005). Furthermore, in 2003 Griffith using PET and fMRI reported directly that lateral part of Heschl's gyrus, adjacent to the primary auditory area, is critical for the pitch perception. He suggested that there is a representation in this area that correlates with the perception of pitch rather than a simple mapping of physical stimulus characteristic and called this area a 'pitch center'.

It was also reported the occurrence of the category specific activation in animal auditory cortex. For example, studies undertaken in primary auditory cortex of awake macaque monkeys (Steinschneider et al., 1994, 1995) and of anesthetized cats (Eggermont, 1995) have suggested a mechanism by which the voice onset time (VOT) phonetic parameter is encoded rapidly in a categorical manner. Consonant-vowel syllables with short VOTs evoked short-latency responses primarily time-locked to consonant release alone. In contrast, consonant-vowel syllables with longer VOTs evoked responses at the same cortical sites time-locked to both consonant release and voicing onset. Hence categorical perception of consonants varying in their VOT could be based partially on temporal encoding mechanisms within primary auditory cortex. The occurrence of two transient response bursts time-locked to both consonant release and voicing onset would signal an unvoiced stop consonant, whereas voiced stop consonants would be represented by a single response time-locked only to consonant release.

More direct evidences originate from studies that have related categorical perception to neuronal activity in behaving animals. Ohl and colleagues (2001) trained gerbils to categorize rising from falling frequency sweeps. During the task, high-resolution electrocortigrams were registered directly over the auditory cortex. As the animals successfully learned to generalize the discrimination of frequency direction, the spatial patterns of longer latency peaks ('marked states') of the electrocortigram changed in a category-specific way.

Therefore the goal of the second part of present study is further analysis of the category-specific activation in the firing of auditory cortex neurons and of the contribution of auditory cortex to the categorization of acoustic signals.

3.2. Methods

3.2.1. Surgery

After completion of the training phase, a head holder was surgically implanted into the monkeys' skull to allow atraumatic head fixation. After retraining with head restrain, a recording chamber implant operation was performed over the left auditory cortex. All surgical procedures were performed under deep general anesthesia with a mixture of ketamine HC1 (2mg/kg) and xylazine (5 mg/kg), followed by a full course of antibiotic and analgetic treatment. Experiments were approved by the authority for animal care and ethics of the federal state of Sachsen Anchalt (number 43.2-42502/2-253 IfN) and confirmed to the rules for animal experimentation if the European Communities Council Directive (86/609/EEC).

3.2.2. Electrophysiological recording

For electrophysiological recordings, a seven- and a sixteen-electrode systems (Thomas Recording) were used. Electrodes were arranged in a circle (in seven-electrode system) or in a 4x4 square (in sixteen-electrode system) and laterally separated from one another by 305 µm. In both cases electrodes could be advanced independently from one another in z-direction. Relatively to the dorsoventral plane electrodes were oriented at an angle of ~ 40 degrees (for monkey F) or of ~30 degrees (for monkey B) and they either penetrated auditory cortex directly or after a traverse of parietal cortex. Recordings were made from a region extending 7 mm in mediolateral direction in monkey B and

6 mm in monkey F and 7 mm in caudomedial direction in monkey B and 8 mm in monkey F, including primary auditory cortex in both monkeys.

After preampflification, the signals from each electrode were amplified and filtered to split them the action potentials (0.5-5 kHz). All data were recorded onto 32-channel analog-to-digital data acquisition systems BrainWave (DataWave Technologies, Minneapolis, MN) or Alpha-Map (Alpha-Omega, Grapeland, TX). On each electrode, the action potentials of a few neurons (multiunit) were discriminate using the built-in spike detection tools of the data acquisition systems (threshold crossing and spike duration). The time stamp and the waveform of each action potential were stored with a sampling rate of 20.833 kHz or 50.000 kHz.

3.2.3. Waveform sorting

From individual multiunit records, the action potentials of a single unit were extracted off-line with a template-matching algorithm (Schmidt, 1984). The waveforms were also used to delete events with artifactual waveforms in multiunit records. The template was created by selecting a number of visually similar and large spike shapes and calculating the average waveform. Subsequently, the waveforms of all events in a multiunit record were cross-correlated with the template, and those waveforms were considered to be generated by the same neuron whose normalized cross-correlation maximum was >0.9. This separation was followed by verifying that there were no first-order interspike intervals <1.5 ms, e.g., smaller that the refractory period of single units in the cortex. The firing of single units, therefore, is included in the firing of some of the multiunits shown below in Results.

The events with artifactual waveforms were also searched in individual multiunit records to extract templates of them and then to detect and delete all events that were attributable to electrical interferences. This procedure was particularly effective in ‘raw’ records in which many events were observed immediately after the switching of the magnetic valve.

3.2.4. Behaving procedure and stimuli

The general layout of the experiment was the same as it was described above. A trial started with the illumination of the LED cue-light. This cue indicated that, within the following 3 s, the monkeys could grasp and hold a touch bar. Once they did so, a tone sequence was triggered 2.22 s later. The first three tones in the sequence had always the same frequency. They were followed by three tones of lower frequency, either immediately or after three to four intermittent tones of higher frequency. For the electrophysiological study the set of sequences was modified and will be described in details below. The monkeys were immediately rewarded with ~ 0.2 ml water when they released the touch bar 240-1240 ms after the onset of the first tone of lower frequency. After bar release, the cue-light was turned off and the tone sequence was stopped, which could happen either during a 200-ms tone or during one of the

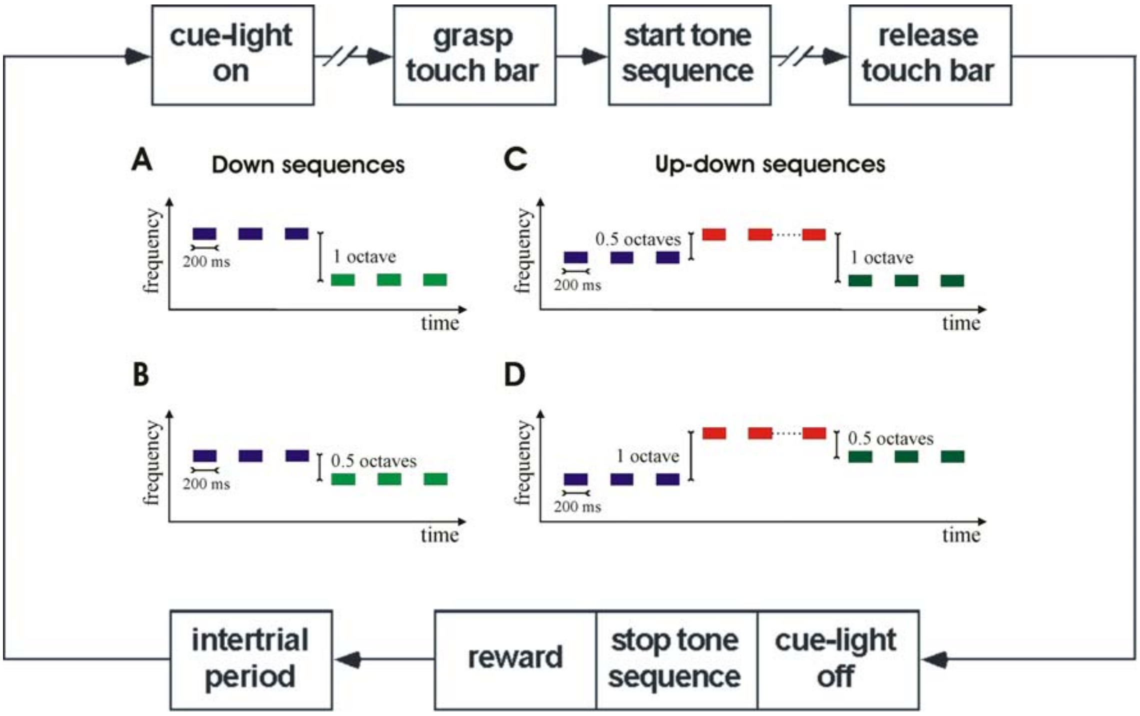

Figure 4: Visual representation of the behaving paradigm. At the beginning of a trial a cue-light was turned on. This was the signal for the monkeys to make contact with a touch bar and to hold it for at least 2.22 s. After this period a tone sequence was started and the monkeys had to release the touch bar upon occurrence of the falling frequency contour in the sequence. If they did so 240-1240 ms after this go-event, the tone sequence was terminated, the cue-light was extinguished, and a water reward was delivered to the monkeys. Bar releases outside the 1000-ms response period resulted in an immediate termination of the tone sequence, the extinction of the cue-light, and a 7-s time-out. The next trial started 6 sec after reward delivery or after the time-out. **(A, B)** Down sequences. The first three tones of the same frequency were followed by three tones of lower frequency. The frequency interval was either 1 octave or 0.5 octaves. **(C, D)** Up-down sequences. The first three tones of the same frequency were followed by three or four tones of higher frequency and then by 3 tones of lower frequency. The frequency interval was either 0.5 octaves by first frequency change and 1 octave by second frequency change or 1 octave by first frequency change and 0.5 octaves by second frequency change. The frequency of initial tones in the sequence varied from trial to trial over a range of 4 to 5 octaves with 0.5 octaves frequency step.

200-ms silent intervals. This was followed by a 6-s intertrial period. When the monkeys prematurely released the touch bar before the go-event a 7-s time-out was added to the intertrial period. In case the monkeys did not release the touch bar during the entire tone sequence the cue-light was extinguished after the last of the three low-frequency tones in the sequence and the 7-s time-out was applied.

The set of stimuli was adjusted for the requirements of neurophysiological experiment. The first modification was an increase in the number of frequency step sizes that could occur within the tone sequences. To be able to get reliable estimates of the responses to different exemplars of frequency contours it was necessary to present individual exemplars several times in the tone sequences. Therefore the number of frequency steps was limited to two, namely 0.5 octaves and 1 octave. The second modification was a reduction of the number of frequencies used in the tone sequences to 12 in monkey F and to 11 in monkey B without restricting the total range of the trial-to-trial frequency variability. The frequencies were separated by 0.5 octaves and the lowest frequency varied between 300 and 500 Hz in different recording sessions, depending on the spectral selectivity of the neurons under investigation. These two modifications resulted in a reduction of the total number of tone sequences from infinite to 31 in monkey F and 29 in monkey B. The set of sequences could be grouped into four types (Fig. 4). Two types concerned to down sequences with only flat and falling frequency contours. They consisted of initially three tones of the same frequency, which were followed by up to four tones with a frequency 0.5 octaves or 1 octave below that of the initial tones. The other two types of sequences concerned to up-down sequences with flat, falling and rising frequency contours. In these sequences, the frequency increased by 0.5 octaves after the initial tones and then decreased by 1 octave, or the frequency first increased by 1 octave and then decreased by 0.5 octaves. Sequences with other combinations of the size of rising and falling frequency steps were not used to minimize the number of different sequences.

The entire set of sequences was played in pseudorandom order to the monkeys. The probability of the occurrence of down and up-down sequences in individual trials was varied between 20 and 80 % by the experimenters interactively to counteract the monkeys' momentary preference during some recording sessions to respond early or late after sequence begin. During individual behavioral sessions monkeys performed up to 628 trials, such that each tone sequence was presented up to 20 times. Mean number of correct trials was 454 (± 128).

Tones in the sequences had a duration of 200 ms with a 10-ms cosine-squared rise/fall and were separated by a silent interval of 200 ms. They were presented at a sound pressure level of ~ 60 dB SPL. The frequency of initial tone in the sequence was varied from trial to trial over a range of 4 to 5 octaves with 0.5 octaves frequency step.

For electrophysiology, acoustic search stimuli (pure tones, noise bursts and frequency sweeps) were generated with a waveform generator (Tucker-Davis Technologies). For a quantitative assessment of the best frequency (frequency of the tone that elicited the maximal number of spikes) and spectral bandwidth of a unit, a random sequence of pure tones was presented. This sequence consisted of 400 tones of 40 different frequencies, usually covering a range of approximately eight octaves (e.g., 0.125-32 kHz) in equal logarithmic steps. Tones had the same duration, envelope, and level as those used for the auditory categorization task. Intertone intervals were 980 ms. Generally these tones were presented at the end of recording session.

3.2.5. Data analysis

To quantitatively determine the early responses of single and multiunits on various tones in the tone sequences, from the neuronal discharges were calculated peri event time histograms (PETHs) with a bin size of 20 ms and a shift window of 1 ms. They were triggered on the onset of each individual tone in the sequence and were calculated separately for different tone frequencies. Only units that met the 4 following criteria were further analyzed. The first criterion was that at least one of the PETHs triggered on the tone onset had to contain a bin with a spike rate that was significantly greater than the spike rate before sequence onset (measured in the 200-ms period immediately before sequence onset; Wilcoxon-tests, $p < 0.05$; most were < 0.01). The bin with the maximal spike rate in a PETH was taken as a measure of the magnitude of the response evoked by individual tones. The center of this bin was used as a measure of the latency of this response. The second criterion for the inclusion of a unit was that its firing rate did not change slowly during the tone sequence. In this case, responses to tones occurring late in the sequence could be different from responses to tones occurring early, which could confound tests of the responses to tones at different positions in the sequence. Such units were identified by exhibiting > 5 consecutive bins in the PETHs that were < 3 standard deviations below the mean pre-sequence firing, 200 ms before sequence onset. Note that the firing activity of these units was analyzed relatively the other features of the behaving procedure and will be described in detail below. The third criterion for the inclusion of a unit was that the response if this unit had to be stationary because the responses on tones from different trials were compared. The forth criterion was that the monkeys performed a number of trials that was sufficient for data analyses (at least 3 presentations of tone of the same frequency for each of the analyzed conditions was necessary). Note that only trials with correct responses were included in this data set. The responses to the pure-tones sequence used by passive stimulation were calculated on the same way as the excitatory responses to the tone sequences using by training.

The selected units were further analyzed for their frequency contour selectivity. It was previously shown that responses of auditory cortex neurons to two-tone sequences are, in part, determined by the frequencies of these tones (Calford and Semple, 1995; Brosch and Schreiner, 1997, 2000; Brosch et al., 1999). To take this into account only about two third of the tone sequences that were presented during the categorization task were used for the characterization of the frequency contour selectivity of units. This was because for each target tone, the sequences were needed in which the preceding tone was 1 or 0.5 octaves lower in frequency (falling frequency contours), had the same frequency (flat frequency contours), or was 0.5 or 1 octaves higher in frequency (rising frequency contours). For each of these stimulus classes exemplars with the same frequency range of the second tone were selected. These exemplars were contained in 19 sequences that were presented during the categorization task. The selected sequences are highlighted with colors in Fig. 5 and comprised of 8 falling, 8 rising, and 4 flat frequency contours. The discarded sequences were essential for the auditory categorization task to make sure that the monkeys actually attended to frequency contours and not to other features of the sequences. Data analyses were performed only on units in which each of the 20 exemplars of frequency contours occurred at least 3 times in the trials with correct behavioral response. The second criterion was that at least one of the PETHs triggered on the tone onset had to contain a bin with a

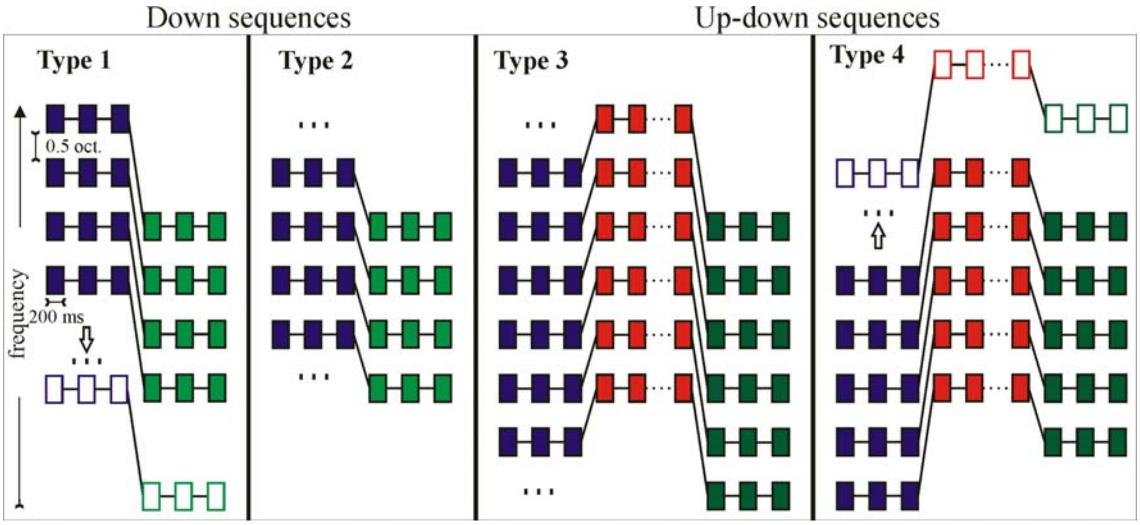

Figure 5: Tone sequences used for the categorization task. The colored rectangles denote the 19 tone sequences that were used to test the frequency contours selectivity of auditory cortex neurons. These sequences are comprised of 8 falling, 8 rising, and 4 flat frequency contours covered 1.5 octaves frequency range. The colorless rectangles denote the tone sequences that were added to the stimuli set to mitigate for monkeys the absolute frequency cues. The frequency of initial tones in the sequence varied from trial to trial over a range of 4 to 5 octaves with 0.5 octaves frequency step. Duration of tones as well as intertone intervals were 200 ms.

spike rate that was significantly greater than the spike rate before sequence onset.

For each stimulus class the response was taken from the bin in the pooled PETH with the maximal spike rate. A unit was defined as frequency contour selective if the pooled response to the second tone of all exemplars of frequency contours of the same class had to be significantly larger than the pooled responses to each of the two other tow classes of frequency contours (Wilcoxon-tests, $p < 0.05$). To avoid the influence of the absolute firing rate of unit a contrast index between falling and rising frequency contours were calculated, in which the difference between the average responses to the falling and rising frequency contours were divided by their sum.

Aside from the neuronal responses to the tone sequence the neuronal firing were also analyzed relatively to the various events of the behavioral procedure. For each record site were calculated PETHs with a bin size of 20 ms, which were referenced to individual events of the behavioral procedure (onset of the cue-light, grasping and releasing the touch bar). First in a PETH time windows of interest were identified in which the firing was > 3 standard deviations above or below the baseline firing measured in the period of 1800 ms before light onset. Then the firing in the time window of interest was compared to the firing in a time window of the same duration taken from baseline. The statistical tests revealed 97.4% of the 2447 time windows of interest (Wilcoxon-tests, $p < 0.05$, most were < 0.01). Note that in the data set were included only units which responded to the tones during the performance of the behavioral task and/or during the passive stimulation and units those were recorded more ventral and less than 1 mm in the supratemporal plane from a site with an auditory response. These criteria guaranteed that only recordings from the auditory cortex entered this analysis.

3.3. Results

3.3.1. Auditory events related firing

Data of this part of the present study were from 25 single units and 94 multiunits in monkey B and from 21 single units and 76 multiunits in monkey F, recorded from the left primary and posterior auditory cortex while the monkeys were engaged in categorizing frequency contours. These units responded phasically with short latencies (30.3 ± 9.6 ms) to the onset of the tones in the sequence. Note that 98 single units and 384 multiunits were excluded from the analysis because they did not responded phasically, because their responses were non-stationary, because they slowly modulated their firing rate during the presentation of tone sequences or because the number of trials performed by the monkeys was insufficient. In addition to the neuronal responses to frequency contours, it will be described also how the units responded to other features of the tone sequences because results on the frequency contour selectivity can be appraised only if they are also put into this broader perspective.

3.3.1.1 Examples

Fig. 6 shows an example of a multiunit from auditory cortex whose firing discriminated various features of the tone sequence. The dots in each row of the dot-rastergrams represent the points in time at which the multiunit fired action potentials relative to the beginning of the tone sequences. Individual tones had a duration of 200 ms and were separated by a silent interval of the same duration. Tones are marked by colored rectangles underneath the abscissa. The three initial tones in each sequence had the same frequency.

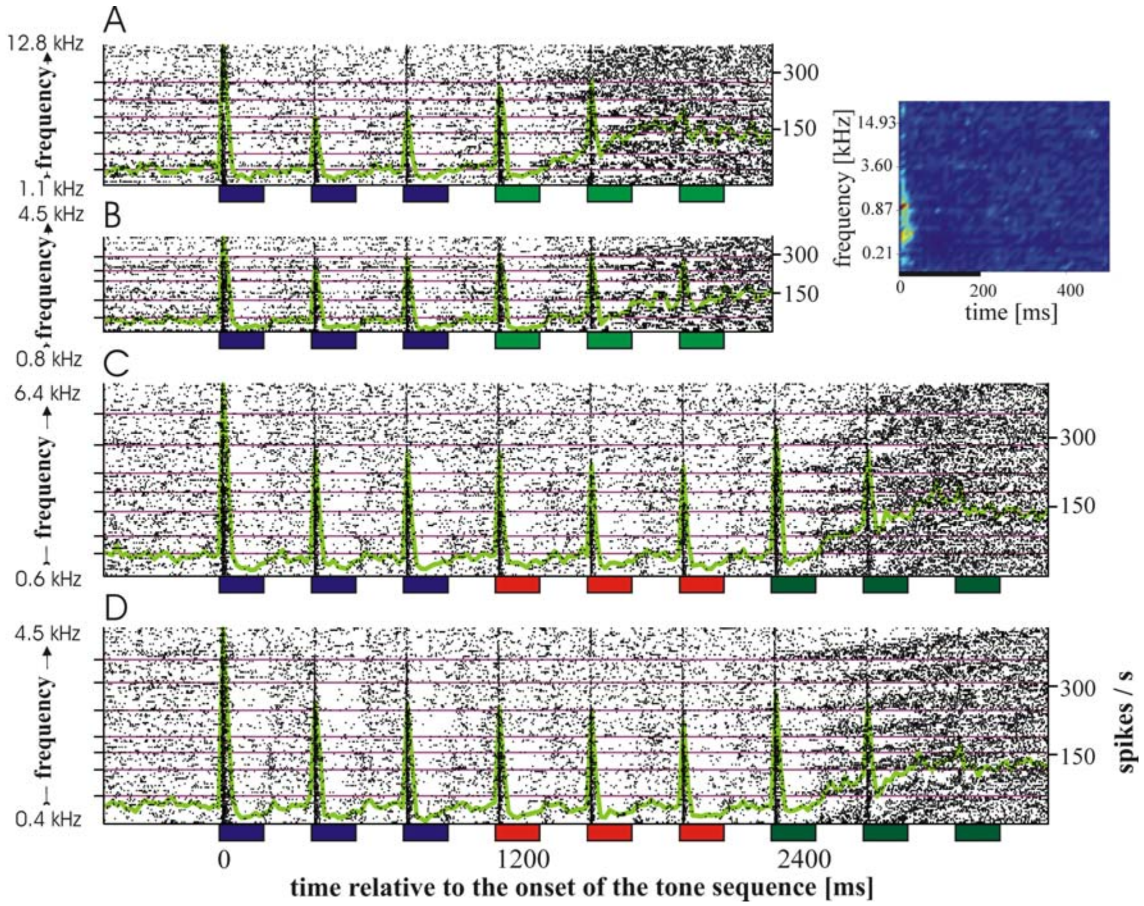

Figure 6: Firing of a multiunit in caudomedial auditory cortex while monkeys were engaged in categorizing frequency contours. Each panel shows a dot rastergram of the neuronal firing and a PETH (green curve) for each of the four sequence types shown in Fig. 5. Each dot represents the point in time when the multiunit fired a spike relative to the first tone in the sequence. The frequency of this tone is indicated on the ordinate. It varied randomly in steps of a multiple integer of 0.5 octaves from trial to trial. Each frequency was repeated 6-21 times. Within each frequency group trials are order according to the reaction time of the monkey i.e. to the time of bar release, from short to long. Colored rectangles denote the tones in the sequence. Note that only 3 tones of higher frequency are plotted because the fourth tone of higher frequency was uncommon in this sessions Starting from the forth tone in the sequence tones had a frequency that was 0.5 octaves (panel A) and 1 octave (panel B) below than that of the initial tones, and that 0.5 octaves (panel C) and 1 octave (panel D) higher. In panels C and D, the frequency changed again with the seventh tone in the sequence. The PETHs were calculated by integrating the responses to all tone sequences, irrespective of the frequency of the tones in the sequence. Inset to the figure shows the frequency response area of this multiunit. It was determined from the firing that was recorded during the presentation of 400 tones at 40 different frequencies before the monkeys categorized frequency contours. Frequency range is indicated on the ordinate. Black bar underneath denotes tone duration. The dark blue color codes the average spike rate during the intertone intervals. Spike rates that are significantly above this rate are plotted with warmer colors.

After them the frequency either stepped down by (A) 1 octave or by (B) 0.5 octaves, or (C) first stepped up by 0.5 octaves and then stepped down by 1 octave, or (D) first stepped up by 1 octave and then stepped down by 0.5 octaves. In each panel, rows were ordered according to the frequency of the initial tone in the sequence, which was varied, in steps of 0.5 octaves, from trial to trial between 0.4 and 12.8 kHz. Trials with the same initial tone, in turn, were ordered according to the time the monkey took to react to the falling frequency contour in the sequence. Only trials with correct responses are depicted.

In each panel of Fig. 6, the vertical bands with an increased dot density indicated that this multiunit responded phasically to the onset of the tones in the sequence. When the responses to different tones were compared to one another it was evident that the largest responses were generally evoked by the first tone in the sequence especially for the sequences those beginning with high-frequency tones. This suggested that the multiunit was most sensitive at the beginning of the sequence. Note that responses to high-frequency tones were also weak when the monkey was passively stimulated outside the categorization task with various pure tones at a rate of 0.85 per sec (inset of Fig. 6). This increased stimulus sensitivity at sequence onset was likely due to the long silent interval of at least 6 sec before the first tone (compared to 200 ms for the tones at later positions in the sequence and to 800 ms in the passive stimulation condition).

Aside from strong responses to the first tone, the multiunit in Fig. 6 exhibited relatively strong responses also to other tones in the sequence, namely to tones that were preceded by a tone of different frequency. These were the responses to the forth tone and, for late-down sequences, also the responses to the seventh tone. By contrast, the responses to the second and third tone and, for late-down-sequences, the responses to the fifth and sixth tone were comparably weak. As these tones were preceded by a tone of the same frequency this indicated that the multiunit was sensitive to changes of the tone frequency in the sequence. The magnitude of the neuronal response to the forth and seventh tone in the sequence even discriminated the type of frequency contour and the size of the frequency interval between consecutive tones. This was hard to be obtained from Fig. 6 because the panels were not arranged according to the size and direction of frequency steps.

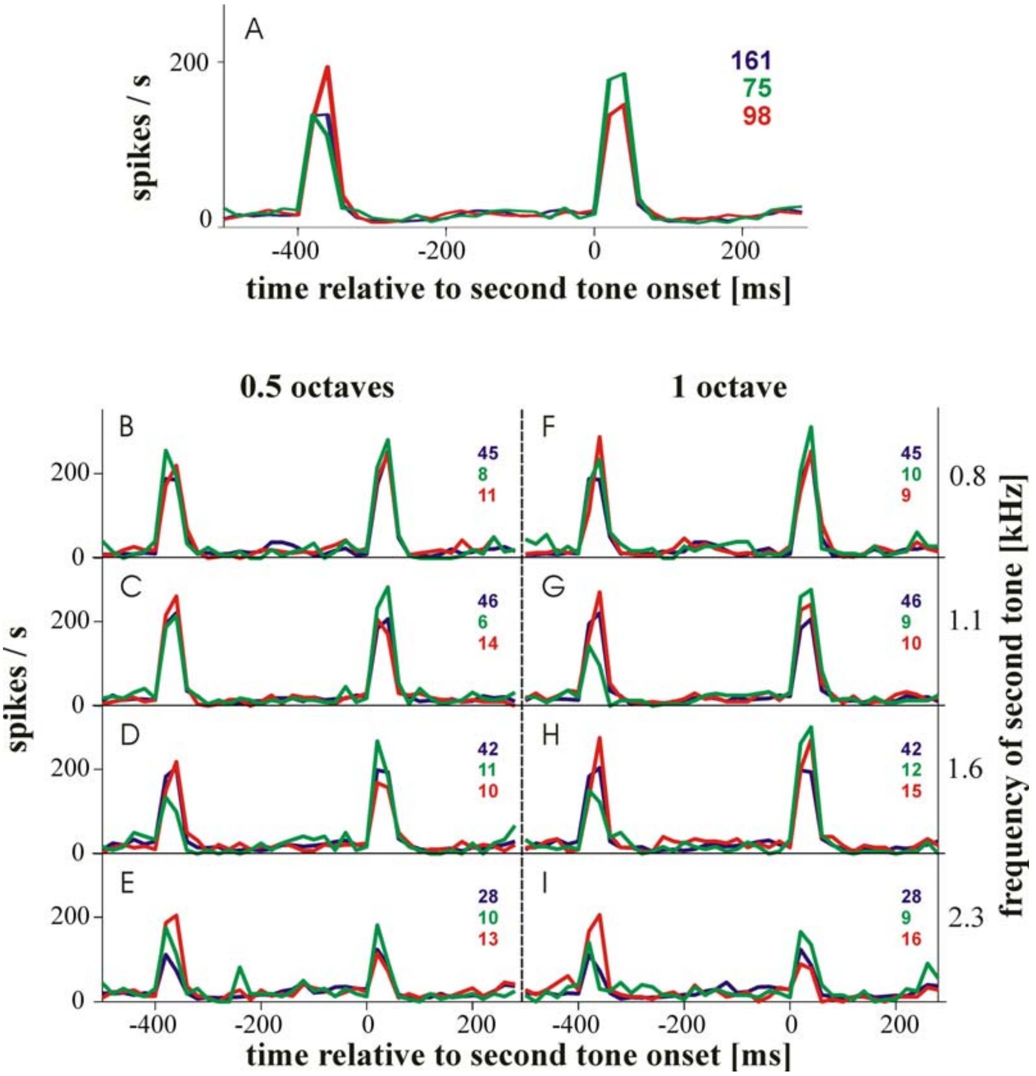

Figure 7: Frequency contour specific neuronal responses in caudomedial auditory cortex. A: PETHs showing the responses to falling (green), rising (red), and flat (blue) frequency contours, relative to the onset of the second tone of a frequency step. Numbers in the upper right corner give the number of tone pairs that were used for the calculation of the PETH. Each PETH was calculated by averaging all corresponding PETHs shown in panels B to I. These panels show the responses to specific exemplars of frequency steps, separately for second tones with different frequencies (rows) and for frequency differences between adjacent tones of 0.5 octaves (left column) and 1 octave (right column). For a better comparison, responses to flat contours (blue) are included in both columns. Note that in each panel, the first tone had a frequency that was different for different frequency contours (0.5 or 1 octave higher as frequency of the second tone for falling contours, 0.5 or 1 octave lower for rising contours and the same for flat contours). Responses to second tone of the same frequency were always largest for falling steps (green curve).

The frequency contour sensitivity became more obvious when neuronal responses were displayed in a different format, namely as PETHs and separately for different exemplars of falling, rising, and flat frequency contours. Fig. 7 shows a multiunit and Fig. 8 a single unit that both responded more

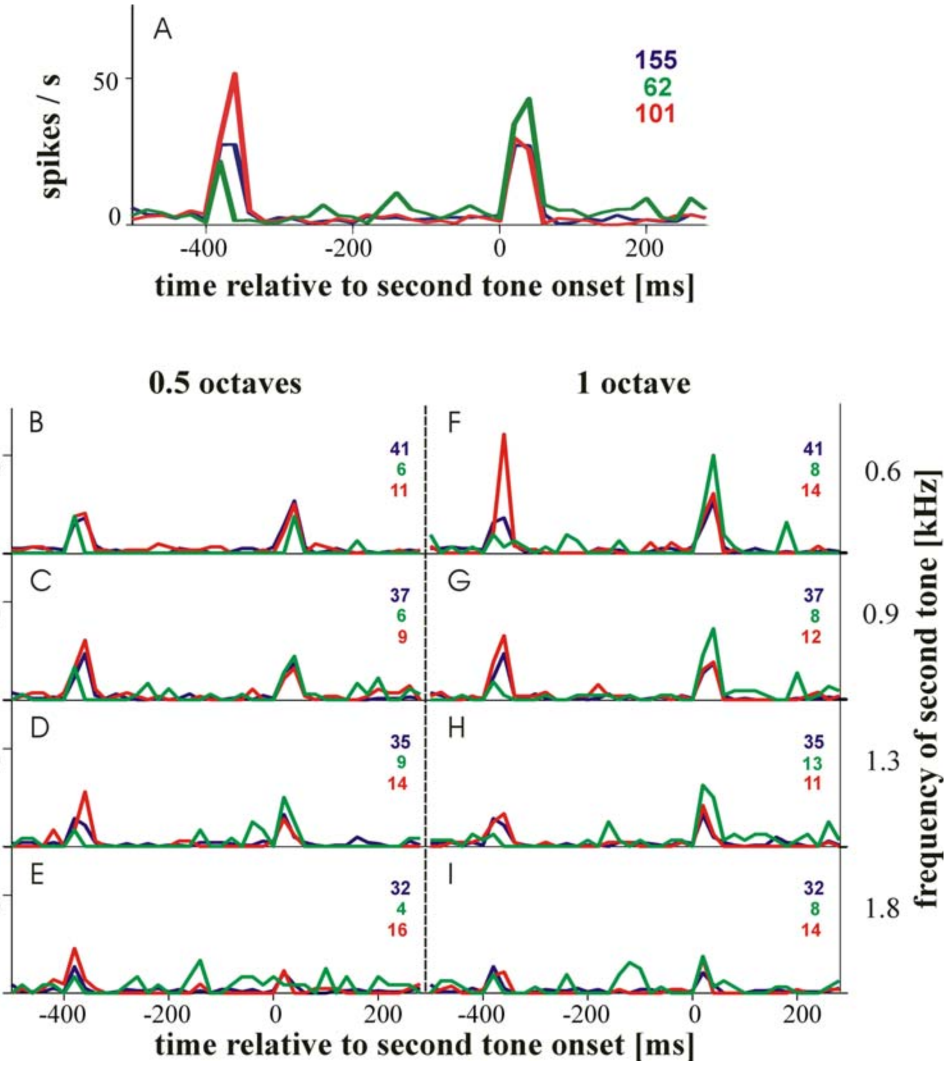

Figure 8: Frequency contour specific single unit recorded from the auditory core field. Same conventions as in Figure 7.

strongly to falling frequency contours than to rising and to flat frequency contours (Wilcoxon-tests, $p < 0.05$). The figures also show responses to individual exemplars of frequency contours, separately for exemplars with different second tones and for step sizes of 0, 0.5, and 1 octave. Note that not for all units the frequency contour selectivity was perfect, i.e. the strongest responses were not all evoked by exemplars of falling frequency contours but some of them were evoked by exemplars of other frequency contours. The reason for this was that the responses also depended on the frequency of the second tone and on the frequency step.

Another property of the multiunit shown in Fig. 6 was its increase in firing towards the end of the sequence. This increase occurred at variable times in different sequences and was synchronized to the release of the monkey's hand from the touch bar. This type of neuronal firing will be described in detail below.

3.3.1.2 Population results

The response properties of the examples shown in Figures 6-8 were also seen in other single and multiunit recordings. There were no systematic and qualitative differences of the response properties between recordings from primary and posterior auditory cortex and between the two monkeys, except that response magnitudes were generally greater in primary auditory cortex. Therefore data were combined for the presentation of the population results.

Adaptation of the responses during the tone sequence

Neurons typically responded most vigorously to the first tone in the sequence. This was the case for 111 multi and 35 single units that were examined in the down sequence condition and for 147 multi and 37 single units that were

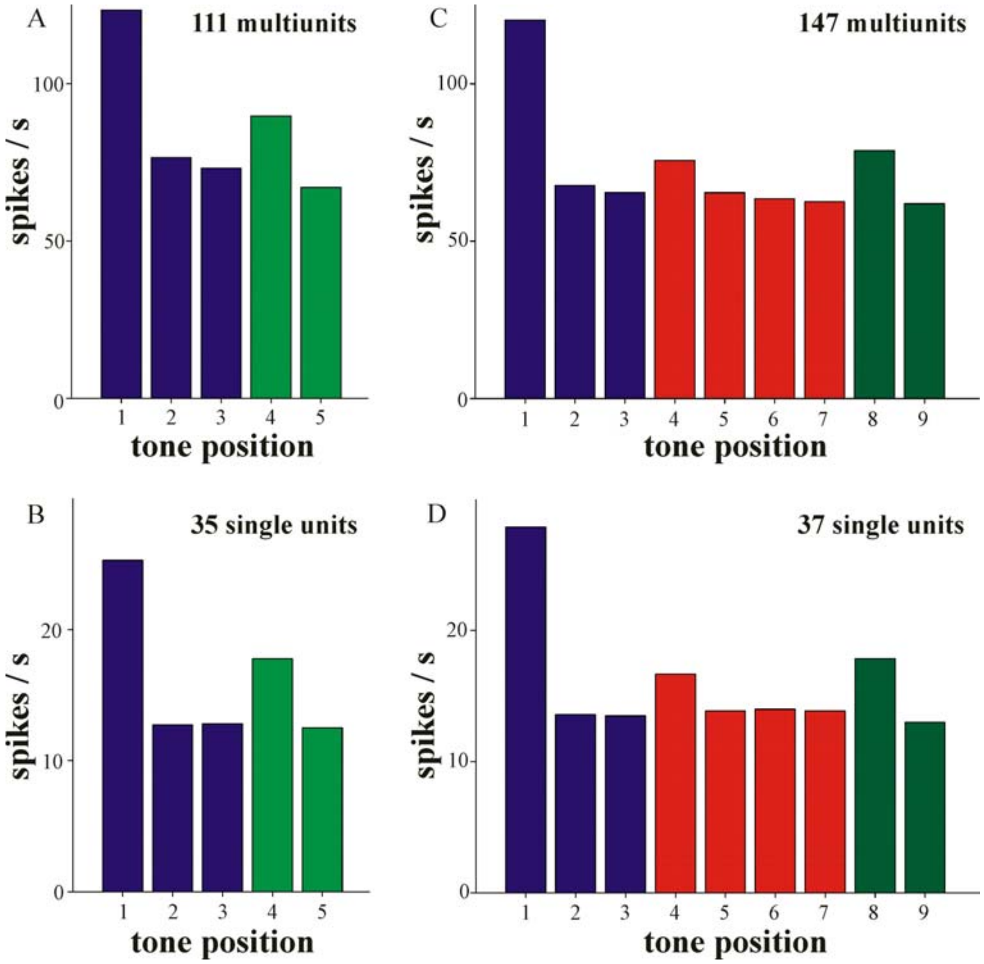

Figure 9: Population responses to tones at different positions in down (A, B) and up-down (C, D) sequences for multi and single units. All pairwise statistical comparisons were significant (Wilcoxon-tests, $p < 0.05$), except for the comparison between the sixth and seventh tone in panels C and D, second and third in panel B and D, and fifth and sixth in panel D. The same frequency range was used for all tones in the sequence. Numbers in upper right corner give the number of units.

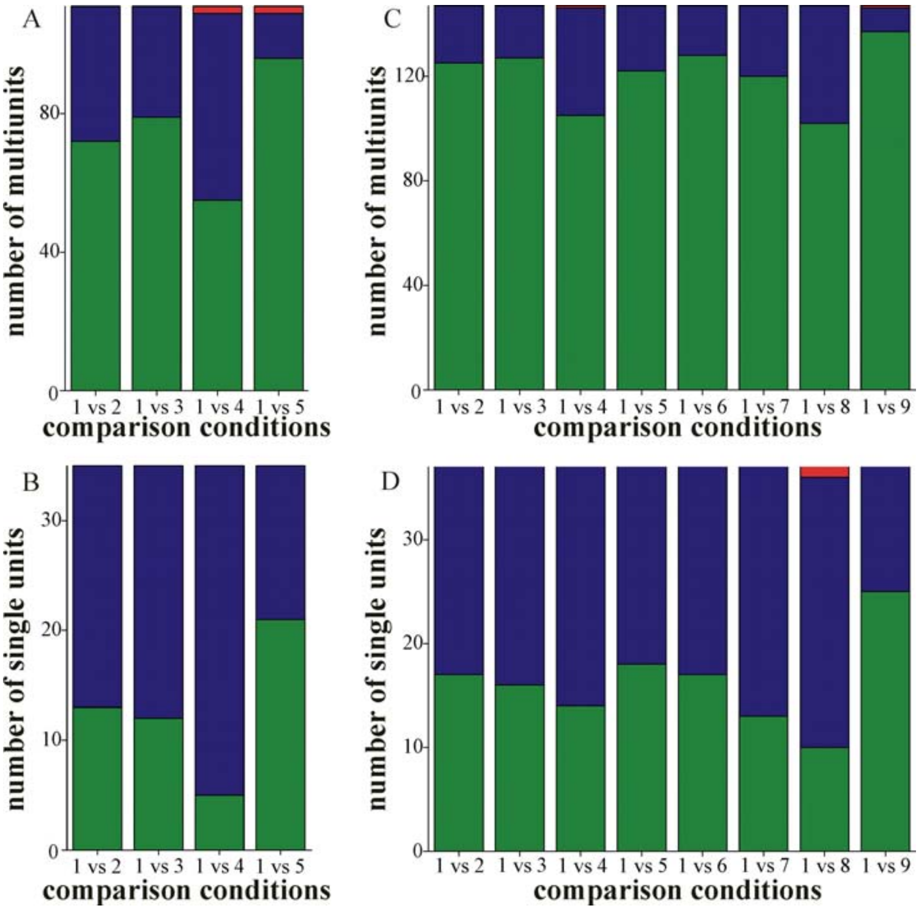

Figure 10: Pairwise comparisons of the responses of multiunits (A, C) and single units (B, D) to the first tone in the sequence with responses to consecutive tones for down (A, B) and up-down sequences (C, D). Green bars give the number of units that responded more strongly to the first tone than to later tones. Inverse response relationships are indicated by red bars. Blue bars denote no response differences.

examined in the up-down sequence condition. The population response to the first tone was significantly greater than the population responses to each of the successive tones in the sequence for both multi and single units samples (pairwise Wilcoxon sign rank tests, all tests with $p < 0.05$; Fig. 9). The difference between the responses to the first and the second tone in the sequence ranged between 45 and 55 % in down and up-down sequences, respectively. For this analysis the same frequency range (3 to 4 octaves) was considered for the calculation of the responses to tones at different positions in the sequence.

The relative enhancement of the response to the first tone was also present in individual multi and single units (Fig. 10). In 73% of multiunits (67/76 by monkey F, 73/94 by monkey B), responses to the first tone were evoked by presentation of the initial tone of the sequence also in the case if presentation of tone with the same frequency didn’t elicit responses for some of the following tones in the sequence, i.e. for the first tone, the frequency range that could elicit a response was broader (Wilcoxon-tests, all tests with $p < 0.05$). For down sequences, 48 of 111 multiunits (43%) and 2 of 35 single units (5%) responded more strongly to the first tone than to any other tone in the sequence. For up-down sequences, this was seen in 84 of 147 multiunits (57%) and in 9 of 37 single units (24%). The percentages were higher for up-down sequences because of the training procedure. Generally more down than up-down sequences were presented during one training session because otherwise monkeys exhibited a tendency to respond to the first frequency change and did not wait for the second frequency change. Therefore less trials with down sequences than with up-down sequences could be examined. On the other hand the frequency ranges of analyzed down and up-down sequences were partly different which could also leads to the different ratios.

Most of the remaining units also responded most strongly to the first tone but not all statistical pairwise tests passed the significance level of $p < 0.05$. Only in a few units, the response to the first tone was significantly weaker than one or more of the responses to the following tones. This was due both to strong responses to the first tone in the sequence and to responses elicited by a wider frequency range. The increased frequency sensitivity at sequence onset was also found when the frequency range that elicited a response to the first tone in the sequence was compared with the frequency range that elicited a response in the passive listening condition, outside the auditory discrimination task.

The most parsimonious explanation for the enhanced sensitivity at the beginning of the sequence was the duration of the silent interval before that tone. The intertrial interval was at least 6 sec for the first tone in the sequence, 200 ms for the tones at later positions in the sequence, and 980 ms in the passive listening condition. There was no significant difference between the responses to the first tone in the sequence when the silent interval was 6 sec (correct behavioral response in the previous trial) and when the silent interval was 13 sec (previous trial was incorrect), suggesting that sequential two-tone interactions lasted maximally 6 sec.

Although the neuronal responses to the tones following the first tone in the sequence were generally weaker than the response to the first tone they could be of different magnitude. One factor shaping the response magnitude at later positions was the size of the frequency difference between the tone evoking the

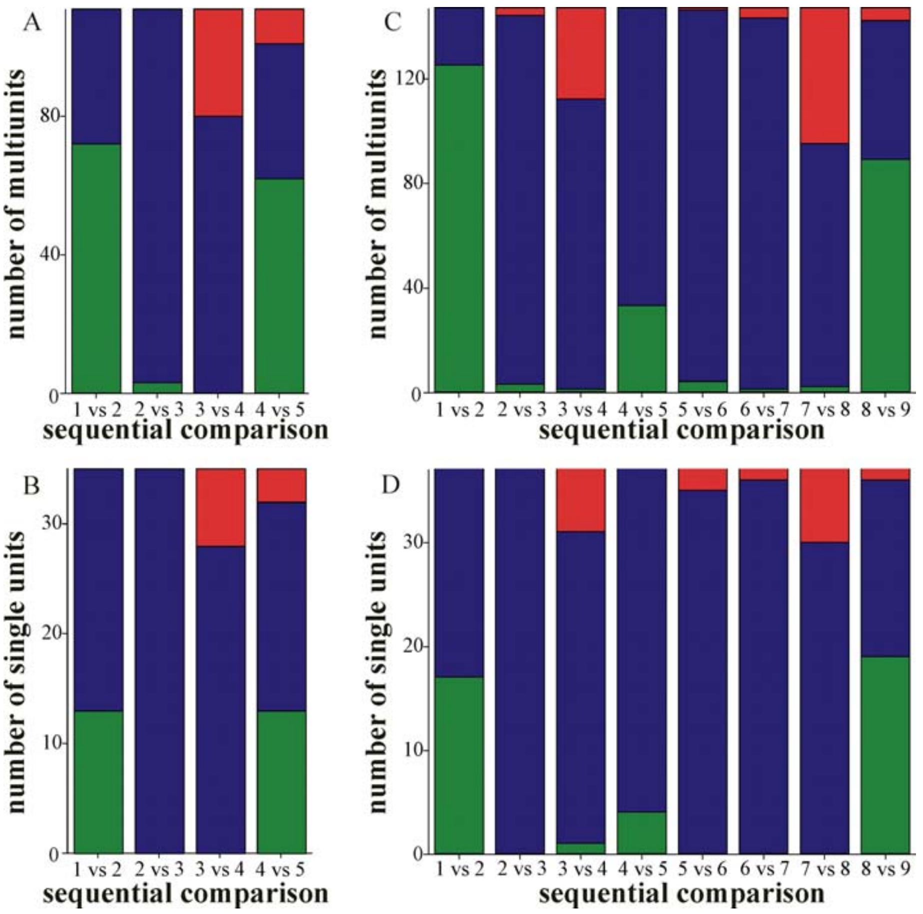

Figure 11: Pairwise comparisons of the responses of multi (A, C) and single units (B, D) to consecutive tones for down (A, B) and up-down sequences (C, D). Green bars give the number of units that responded more strongly to the first of two adjacent tones. Inverse response relationships are indicated by red bars. Blue bars denote no response differences.

response and the preceding tone. Both for single and multiunits, population responses were weak (Fig. 9) and the number of units with differential responses to tones within the sequence was low (Fig. 11) for the second, the third, and the fifth tone in the sequence, i.e., for tones that were preceded by a tone of the same frequency. For up-down sequences, this applied also to the responses to the sixth, seventh and ninth tone. In multiunit sample responses became progressively weaker by second and third repetitions of tone of the same frequency. Although the decrement in responsiveness by second repetition of the same tone was significant (p = 0.03), in the population it was comparably small (1.8%) compared to the response changes that occurred from the first to the second presentation of the same tone, which ranged between 45 and 55%. Population responses to the third repetition of the same tone (namely to the seventh tone in the up-down sequences) were not significantly different from the responses to the previous tone. This suggests marginal higher-order sequential interactions between non-adjacent tones, which was also reflected in the firing of individual single unit and multiunit recordings of which very few exhibited significantly different responses to tones that were repeated more than once. This suggest also that although flat frequency contours were not presented at the same position as falling and rising frequency contours, the response at the third position in the sequence could be used to estimate the response to flat frequency contours at the forth and eighth position.

By contrast to the responses to the tones of the same frequency, responses were of intermediate magnitude when a tone was preceded by a tone of different frequency. This was the case for the responses to the forth tone, and for up-down sequences for the response to the eighth tone. Both the forth and eighth tones elicited a population response that was stronger than the population response to the previous tone (Wilcoxon-tests, all tests with $p < 0.01$). An increment in responsiveness by the change of the tone frequency was 26.43% (19.13-21.87%). Responses to the first tone after the frequency change were significantly stronger than responses to the previous tone in 31 of 111 multiunits (28%) and in 7 of 35 single units (20%) when down sequences were analyzed. For up-down sequences, 35 of 147 multiunits (24%) and 5 of 37 single units (14%) responded more strongly to the tone if its frequency was higher than the frequency of the previous tone and 52 multiunits (35%) and 7 single units (19%) responded more strongly to the tone if its frequency was lower than the frequency of the previous tone.

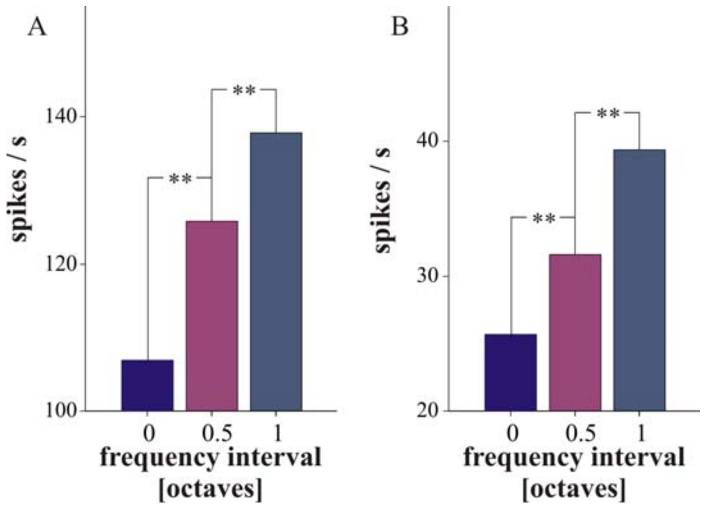

Figure 12: Dependence of the responses of multiunits (A) and single units (B) on the size of frequency steps at a stimulus onset asynchrony of 400 ms.

Generally the response to a given tone in the sequence was the stronger the larger was the size of the frequency interval between this tone and the preceding tone (Fig. 12). Note that for this analysis a smaller number of units was used namely the same set as used for the characterization of frequency contour selectivity (see below). This was intended to avoid the influence of the direction of frequency change and, for statistical reasons, to have at least 3 presentation of tone of the same frequency in all conditions. The response to tone pairs with an 1-octave step was 10 % greater than the response to tone pairs with a 0.5-octave step ($p < 1.5*10^{-7}$) and 29 % greater than the response to tone pairs of the same frequency ($p < 3.3*10^{-11}$). The latter responses, in turn, were 18 % smaller than responses to tone pairs with a 0.5-octave frequency step ($p = 3.1*10^{-11}$). The dependence of the response on the size of the frequency step was also observed in the firing of individual multi and single units where 8 of 59 multiunits (14%) and 2 of 15 single units (13%) responded more strongly after an 1-octave frequency step than after an 0.5-octave frequency step and after a tone of the same frequency. No unit exhibited the opposite response behavior, i.e., responded more strongly after an 0.5-octave frequency step or after a tone of the same frequency. No such differential responses were found for stimuli across the intertrial interval of 6 sec, e.g. between the response to the first tone of a given tone sequence and the response to the last tone of the previous trial which had the same or a different frequency. In all of these tests the same frequency range was used for the calculation of the responses to tones at different positions in the sequence.

This observations extend previous findings on the dependence of neuronal responses on the frequency and time separation of tone pairs in auditory cortex of non-performing animals (Calford and Semple, 1995; Brosch and Schreiner, 1997, 2000; Brosch et al., 1999; Bartlett and Wang, 2005; Brosch and Scheich, submitted) by demonstrating that, during active listening, the response to a tone can depend on the frequency difference between this and the preceding tone even when the onset asynchrony of the two tones is as long as 400 ms.

Frequency contour selectivity

Units in auditory cortex responded more strongly to falling frequency contours than to rising and flat frequency contours (Fig. 13). This was investigated in 4 single units and 27 multiunits from monkey F and in 11 single units and 32 multiunits from monkey B, which had BFs between 0.15 and 23 kHz. These units comprised only a subset of the total of tone-selective units because of statistical reasons it was required that the monkeys correctly responded to falling frequency contours for at least 3 of each of the 20 exemplars of frequency contours in the tone sequences, as outlined earlier. For down

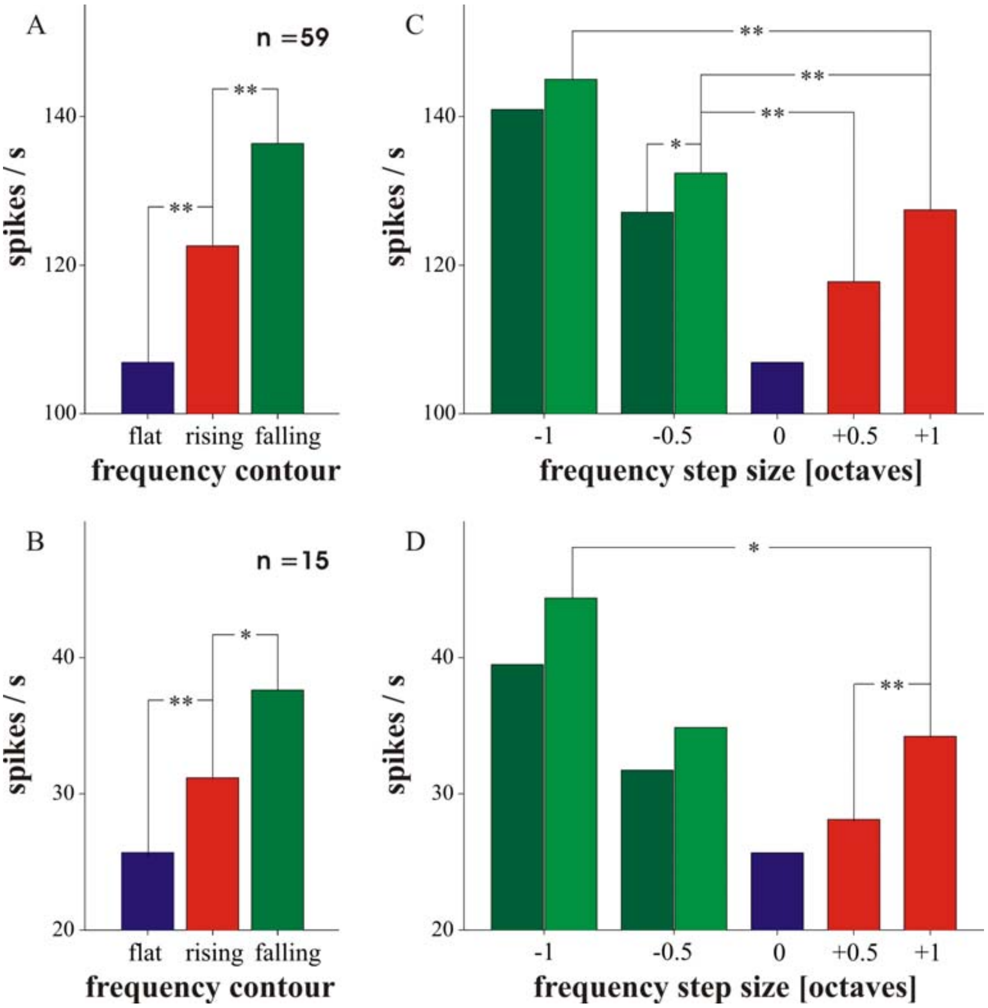

Figure 13: Population results of the frequency contour selectivity in auditory cortex for multiunits (A) and single units (B). C, D: Dependence of population responses on frequency contours, their position in the sequence, and size of frequency steps for multiunits (C) and single units (D). Statistically significant differences between falling and rising frequency contours are marked by asterisks (Wilcoxon-tests, *: $p < 0.05$, **: $p < 0.01$). Differences between frequency steps were significant ($p < 0.05$) but not marked.

population response to falling frequency contours was 13.2 % stronger ($p < 0.001$) than the response to rising frequency contours, both in the multiunit and in the single unit sample. The population response to rising frequency contours was also 14.7 % stronger than the response to flat frequency contours ($p < 0.0001$). For up-down sequences the difference between the population responses to falling and rising frequency contours was less strong (9.3 %) but still significant ($p < 0.0001$). The reason for the difference between down and up-down sequences was that the responses to rising and falling frequency contours were tested at different points in time during the tone sequence. Because some units slightly decrease their firing during the sequence, responses to tone occurring later in the sequences tended to be weaker than those occurring earlier. This phenomenon could be also related to the behavioral procedure. Monkey should discriminate the direction of the frequency change only after the first frequency change. The second frequency change was always downwards, so in this case it was necessary for monkeys just to detect but not to discriminate the change.

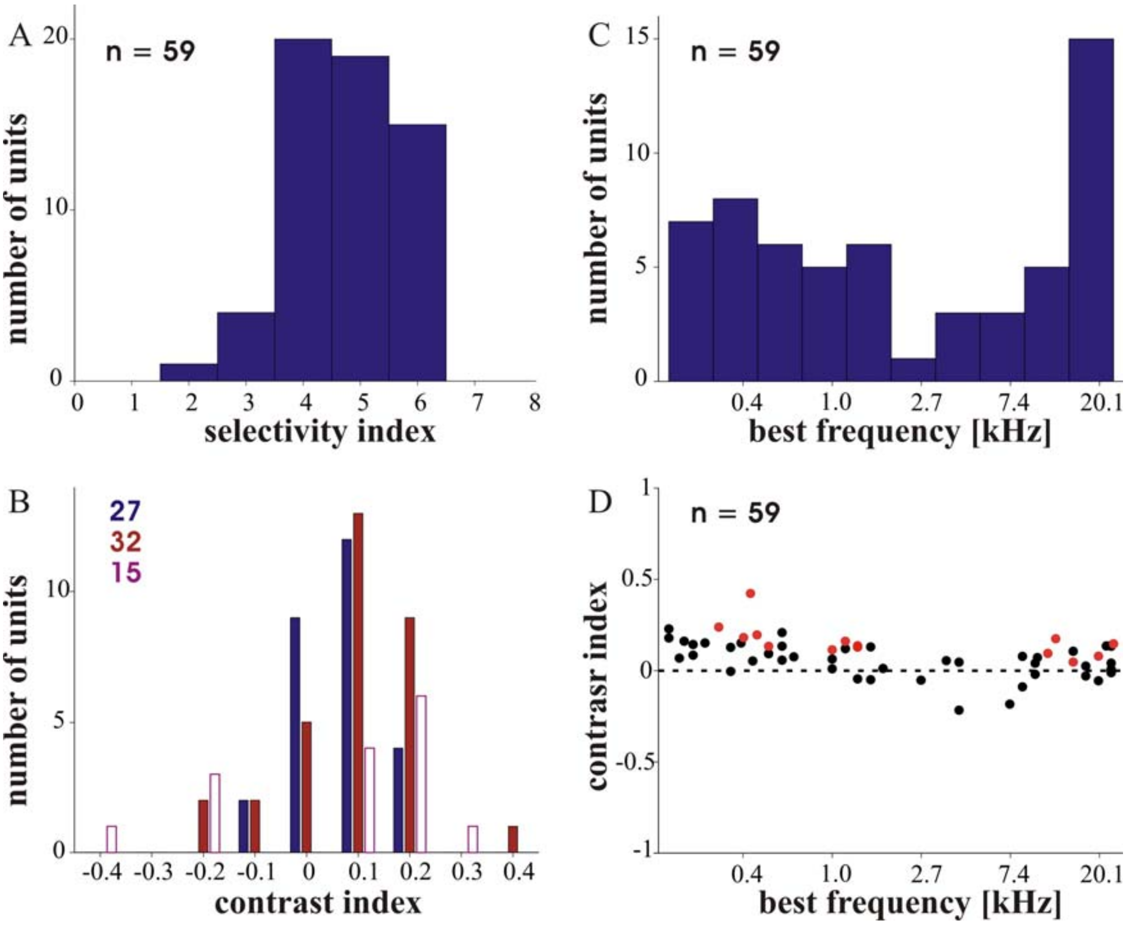

Figure 14: Degree of frequency contour selectivity of neurons in auditory cortex. A: Sensitivity index. Distribution of the number of exemplars of falling frequency contours that were among the 8 maximal responses of multiunits to 20 exemplars of falling, rising, and flat frequency contours. The distribution was significantly different from distributions that would obtained if the responses were independent of the type of frequency difference between adjacent tones ($p < 0.01*10^{-4}$, bootstrap). B: Category tuning. A contrast index was calculated for each unit by dividing the difference of the responses to all exemplars of falling and of rising frequency contours by their sum. The response to falling frequency contours was calculated only from down sequences. Colored bars show the distribution of contrast indexes for multiunits (blue for monkey F, red for monkey B), colorless bars show the distribution of contrast indexes for single units. C: Distribution of the BF’s in multiunit sample. The best frequency of frequency contour sensitive units varied between 0.16 and 23.48 kHz. D: Correlation between frequency contour selectivity and the best frequency of units. Red dots correspond units those responded significantly stronger to the falling frequency contour than to the rising frequency contour (frequency contour selective units), black dots show units with no significant difference between responses to the falling and rising frequency contours.

The preference for falling frequency contours was also found in the responses of individual units, of which 1 of 15 single units (7 %) and 14 of 59 multiunits (23.7%) responded significantly more strongly to falling than to rising frequency contours and to flat frequency contours (Wilcoxon-tests, all tests with $p < 0.05$) but no unit responded significantly more strongly to rising frequency contours.

When the responses to different exemplars of frequency contours were compared it was found among the 8 greatest responses, on average, 5 (± 1) exemplars of falling frequency contours (Fig. 14, A). The distribution of the number of large responses to falling frequency contours was significantly different from the distribution that is expected if units were insensitive to the type of frequency contour ($p < 0.01 * 10^{-4}$, insensitive data sets were created with bootstrap procedures) or if the responses of units were only affected by the size of the frequency step ($p < 0.01$) Another way of assessing the degree of frequency-contour selectivity of multiunits was the contrast index between falling and rising frequency contours (Fig. 14, B). The average contrast index was 0.081 for the multiunit sample (0.073 for monkey F, 0.088 for monkey F, in both cases significant different from 0, Wilcoxon-tests, $p < 0.0001$) and 0.051 for the single unit sample. No correlation was found between the contrast index and the best frequency of unit. The best frequency of frequency contour selective units varied within a wide range of frequencies namely between 0.28 and 23.48 kHz.

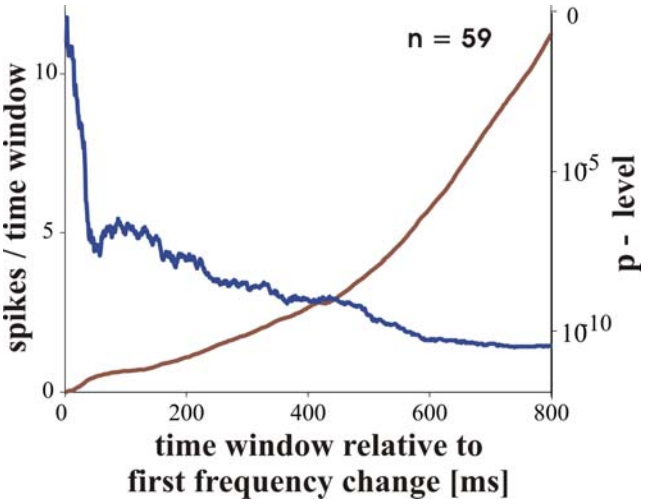

Figure 15: Dependence of the response difference between falling and rising frequency contours on the duration of the analysis window. The analysis window commenced at onset of the second tone of a frequency contour and was progressively increased. The average difference of the response to falling and rising frequency contours is shown in red and was computed from 59 multiunits. Its statistical significance was tested with Wilcoxon signed rank tests and is shown in blue.

To more thoroughly define at what points in time multiunit responses in auditory cortex discriminated between falling and rising frequency contours, an analysis was performed in which the length of the analysis window was progressively increased. As shown in Fig. 15, the firing discriminated the two types of frequency contours at a significance level of $p < 0.05$ when the window was 11 ms. The significance level rapidly dropped to lowest levels within 10 ms. This suggests that earliest parts of the response to the second tone of a frequency step discriminated frequency contours. From Fig. 15 it can also be obtained that the difference between the neuronal firing associated with two types of frequency contours became progressively stronger with the increase of the analysis window's length. This is due to monkeys' task performance. Since only the successful trials were analyzed, in the time period after first frequency change monkeys ether released the touch bar by occurrence of the falling frequency contour or continued to grasp them by occurrence of the rising frequency contour.

Dependence of frequency contour selectivity on task performance

In order to calculate the dependence of neuronal response on the task performance single trials analysis was performed. Four types of trials were compared: successful trials where monkeys released the touch bar after occurrence of the falling contour in down sequences (hits) and didn't release them after occurrence of the rising contour in up-down sequences (correct rejection), unsuccessful trials where monkeys premature released the touch bar in 240-1240 ms after onset of the first tone of higher frequency in up-down sequences (false alarm) and where monkeys continued to grasp the touch bar until the end of the down sequence (miss). Since the number of trials of each type was different they were randomly selected a number of times. Each set of corresponding trials was arranged relative to the response strength and got ranks from 1 to 4. Note that the target frequency and the frequency interval were the same in all four cases.

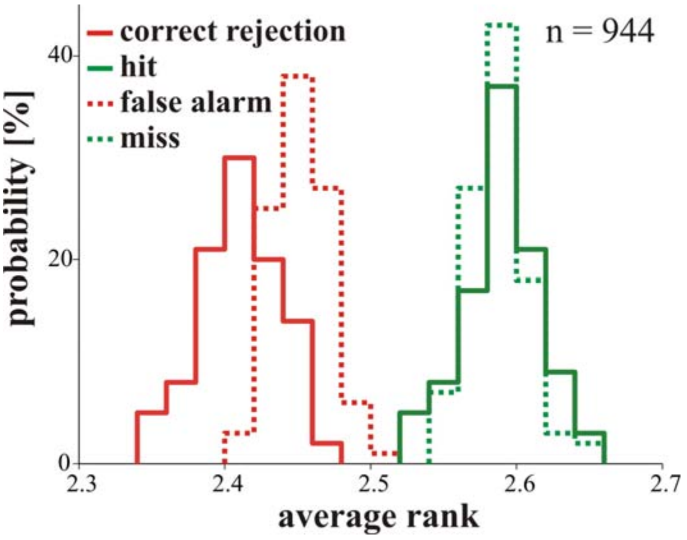

Figure 16: Dependence of neuronal response to the frequency contour on task performance. Average rang was calculated from randomly selected corresponding trials. Frequency contour is color coded. Solid curves show the neuronal firing in successful trials, dash curves show fail trials.

As shown in Fig.16, the neuronal response to the rising frequency contour was stronger in false-alarm trials than in successful trials (Wilcoxon-test, $p < 0.0001$). No difference was found between neuronal responses to falling frequency contours between successful and unsuccessful trials.

3.3.2. Non-auditory events related firing

Data of this part of the present study were from 315 multiunits and 64 single units from the auditory cortex of the two monkeys. These units were examined whether they exhibited the changes in the firing rate synchronized to the non-auditory events that coordinated the auditory categorization task. In the

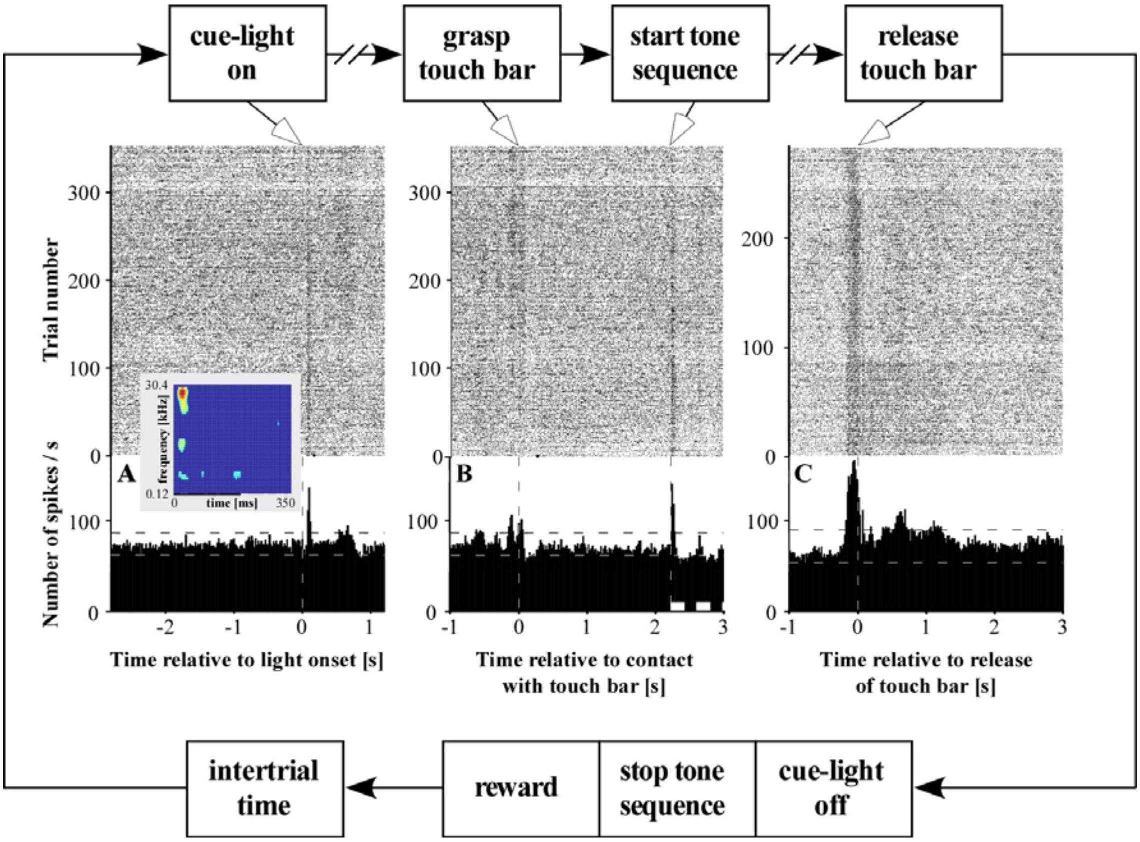

Figure 17: Multiunit in primary auditory cortex that fired action potentials during acoustic and non-acoustic events of an auditory categorization task. Outer flow diagram depicts details of the behavioral task. Inner panels show dots rastergrams and peri event time histograms computed from the neuronal discharges, which were temporally aligned to (A) the onset of the cue-light, (B) the bar touch and the onset of the tone sequence, and (C) the release of the touch bar. Number of trials is given in the upper right corner of each panel. Panel C was computed from trials with correct responses only. Dashed horizontal lines mark three standard deviations above and below average baseline firing, computed from the period of 1800 ms before light onset. In panel A, inset shows the response plane of this unit, computed from the spikes that were recorded during the presentation of pure tones within the frequency range of tones indicated on the ordinate. In panel B, the white bars mark the first and the second tone in the sequence.

following particularly clear examples of these types of firing behavior are shown, and then the group results are present.

3.3.2.1. Examples

Fig. 17 shows examples of a non-auditory event-related firing in a representative multiunit from primary auditory cortex. The three peri event time histograms were triggered on the onset of the cue-light (panel A), on the moment the monkey's hand touched the bar (panel B), and on the moment the monkey released the bar (panel C). Each of the PETHs consists of one or several peaks, which indicates that the neuronal firing was synchronized to these events. The unit fired 80-120 ms after the cue-light went on, 120 ms before until 80 ms after the monkey had contacted the touch bar, and from 160 ms before until 1220 ms after the release of the bar. Figure 23, B also shows the responses to the first two tones of the sequence, which started 2.22 s after bar touch.

Aside from phasical responses to the non-auditory events many units also exhibited slow modulations of their firing. The firing of the multiunit shown in Fig. 18, A slowly decreased ~ 700 ms after the grasp and continued to do so until the onset of the tone sequence (this distinguished slow changes of firing from

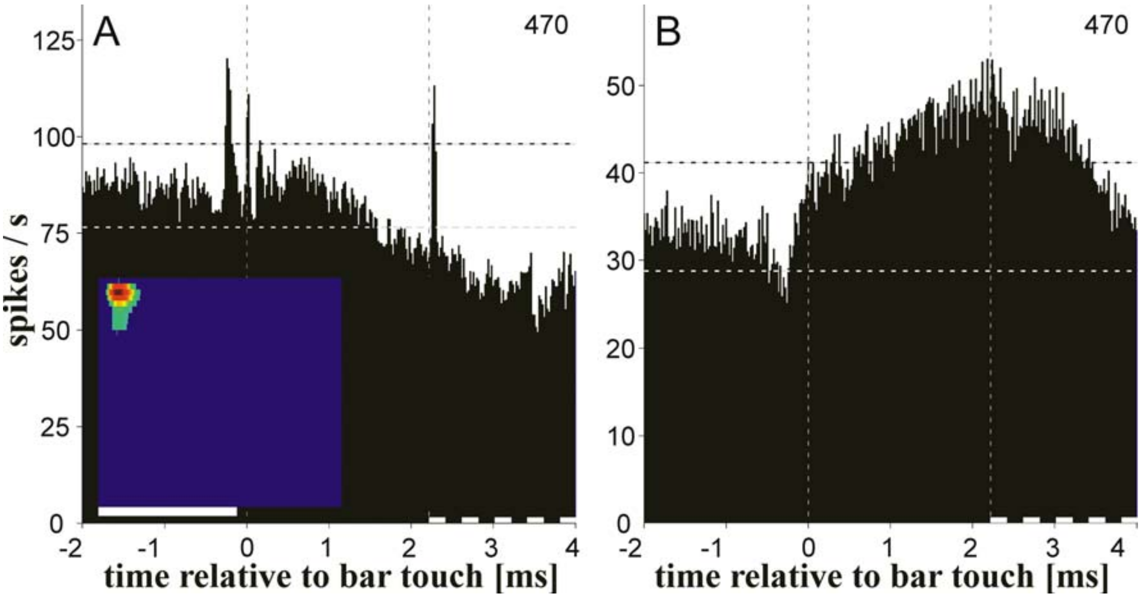

Figure 18: Examples of slow changes in firing rate after bar touch. (A) Multiunit with a slow decrease of firing during the bar holding period. The two narrow peaks around zero indicate transient firing synchronized to bar touch. The single peak 2.22 s after bar touch indicates the response to only the first tone in the sequence. (B) Multiunit with a slow increase of firing during the bar holding period. This unit could not be driven by tones during and after the auditory discrimination task and therefore no response plane is shown. Note that both multiunits were simultaneously recorded from different electrodes.

transient firing). Fig 18, B shows a unit that was recorded simultaneously from the other electrode. Firing of this unit exhibited the opposite behavior, e.g., slow increase after the grasp that reached the significant level ~800 ms before tone sequence was started.

3.3.2.2. Population results

Cue-light related firing

45 (14.3 %) multiunits and one single unit (1.7 %) exhibited firing that was synchronized to the onset of the cue-light. In 39 multiunits and the single unit this firing was significantly stronger than before light onset, whereas the firing decreased in 6 multiunits (Wilcoxon-tests, $p < 0.05$). Increases in firing ranged from 119 to 700 % (median 139 %), decreases from 57 to 75 % (median 70 %). Latencies of cue-light firing ranged between 60 ms and 240 ms after light onset (median latency 100 ms). In all units the increase in firing was transient, with a median duration of 60 ms (range 20-220 ms). These latencies and response durations corresponded well to those in early visual cortical areas (Raiguel et al., 1989). To describe the dynamics of the entire population as a function of time, a recruitment function was computed. Fig. 19 shows the number of units that fired in different time bins relative to the onset of the cue-light. This function reveals that the maximal number of units fired 120 ms after the onset of the cue-light and that cue-light related firing disappeared within 360 ms after light onset.

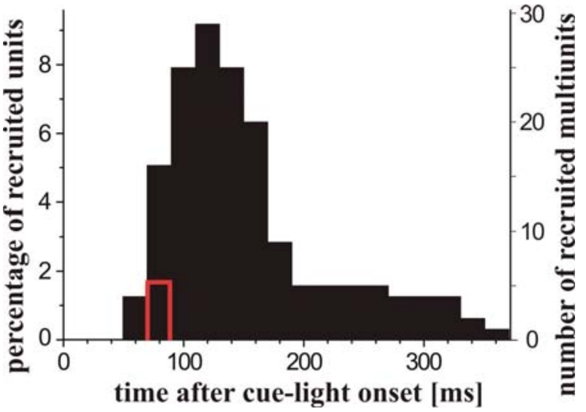

Figure 19: Modulations in neuronal firing after the cue-light onset. Recruitment of units expressed as the percentage or number of units with a significantly modified firing rate in a given bin relative to the onset of the cue-light. Black bars correspond multiunits, red curve shows a single unit.

In 43 of 45 units such cue-light related firing was not observed in error trials, i.e., in trials in which the monkeys did not grasp the touch bar during the 3-s period after light onset (Fig.20). This suggests that, in many neurons, light-related firing was evoked only under specific conditions, such as when the monkey attended to visual stimuli or associated the cue-light with auditory

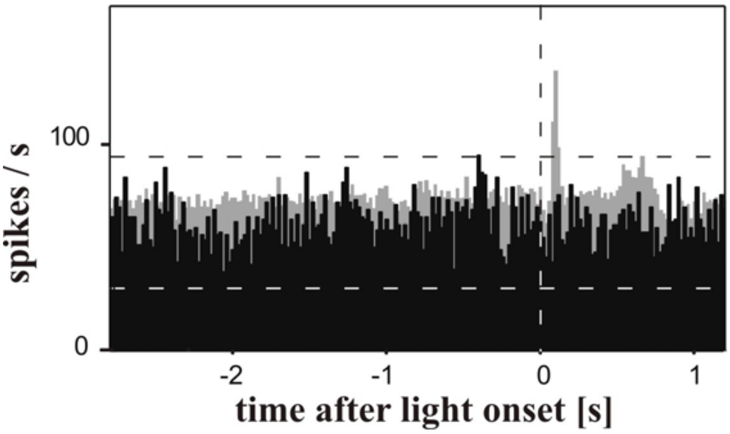

Figure 20: Performance-dependence of cue-light related firing in auditory cortex. The PETH computed from 41 trials in which the monkey did not touch the bar while the cue-light was on is plotted in black. The PETH for the 352 trials in which the monkey made contact with the touch bar after the cue-light was lit is plotted in gray.

processing, or when the visual stimulus fell into the visual receptive field of a cell in auditory cortex.

Grasping-related firing

Firing associated with the grasping of the touch bar was observed in 194 multiunits (61.6 %) and 21 single units (35.0 %) in which the PETH triggered on bar touch consisted of bins significantly different from baseline firing (Wilcoxon test, $p < 0.05$). Like the example shown in Fig. 17, these units transiently increased or decreased their firing for a period of 20-1080 ms (median 100 ms for multiunits and 120 ms for single units). The first unit increased its firing 340 ms before bar touch (Fig. 21). Subsequently the number of firing units increased and reached a maximum 20 ms after the hand had touched the bar. During the following 800 ms of the hold period the number of recruited units with transiently elevated firing gradually decayed. Note that it was not detect any sounds associated with the grasping of the touch bar that might have evoked such firing.

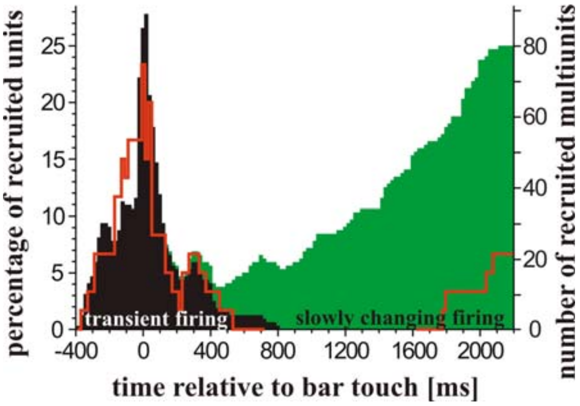

Figure 21: Modulations in neuronal firing after monkeys grasp the touch bar. Recruitment of units expressed as the percentage or number of units with a significantly modified firing rate in a given bin relative to bar touch. Black and green bars correspond multiunits, red curves show single units.

Slow decreases of firing were seen in 37 multiunits (11.8 %) and 4 single units (6.7 %). These slow changes of firing became significant at the earliest 140 ms after the monkeys had touched the bar (Fig. 21). Subsequently, the percentage of multiunits with these properties increased and reached a value of ~ 25 % when the tone sequence began.

The opposite behavior, a slow increase of firing during the hold period, was seen in 40 multiunits (12.7 %). The population response of these units is plotted in Fig. 22. 60% of these units (24/40) exhibited also phasic responses to individual tones of the sequence. Note that all of these units were recorded from the caudomedial field of auditory cortex. No unit from primary auditory cortex exhibited this type of firing. Furthermore, no single unit with this type of firing was found.

After the monkey grasped the touch bar, these multiunits slowly increased their discharge rate. The firing rate reached the level significantly above the baseline firing 670 ms before onset of the tone sequence (200-1000 ms) and continued increasing while the monkey waited for the tone sequence and also after onset of the tone sequence. The latency of the maximal firing rate was 747.5 ms (500-1000 ms) after sequence onset.

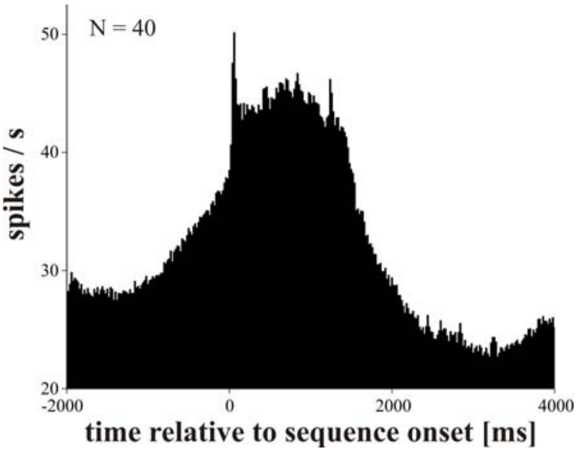

Figure 22: Slow increase in firing during the hold period, observed in 40 multiunits. After monkey grasps the touch bar, firing rate slowly increased and reached the maximum in 500-1000 ms after onset of the tone sequence. Then firing returned to the pretrial level.

For down sequences, the firing returned then quickly to pretrial firing. It occurred in 430 ms (200-600 ms) after the falling frequency contour. For up-down sequences, the reduction in firing occurred over a longer time scale and reached the pretrial level in 830 ms (400-1100 ms) after the rising frequency contour. The latency of the minimal firing was 1060 ms (1000-1300) after the frequency change for down sequences and 1920 ms (1600-2300) after the first frequency change for up-down sequences.

Release-related firing

As described above (see Fig. 6), many units increased their firing rate about end of the tone sequence. In successful trials the tone sequence ended with the bar release following the falling frequency contour. In order to disentangle the firing related to each of the two events four PETHs were calculated. Each PETH was triggered on either the onset of the first tone of lower frequency or the bar release, and they were established separately for trials with different reaction times of the monkey. This is exemplified in panels A and B of Fig. 23, which each show two PETHs of a multiunit whose firing was triggered on onset of the falling frequency contour and on the bar release, respectively. In both panels one PETH was calculated from trials in which the monkey released the touch bar early (reaction time < 600 ms) after occurrence the falling frequency contour, and another PETH in which the release occurred late (reaction time > 800 ms).

In panel A, the PETHs for early and late behavioral responses initially consist of narrow peaks in register. This indicates that, independent of the monkey's reaction time, the transient firing is precisely coupled to the onset of the tones in the sequence. By contrast, the slow increase of firing rate about the end of the tone sequence depends on the monkey's reaction time. A high and broad peak emerged about 500 ms after onset of the lower frequency tones in trials with early releases (red bars) and about 900 ms in trials with late releases (black bars). These observations indicate that the occurrence of this strong firing was related to the variable time of bar release.

Fig. 23, B shows the PETH of this multiunit whose firing in this case was triggered on the bar release. This PETH consists of three narrow peaks around the time of bar release and two broader peaks well after bar release, which are all in register both for trials with early and late releases. Conversely, the multiple

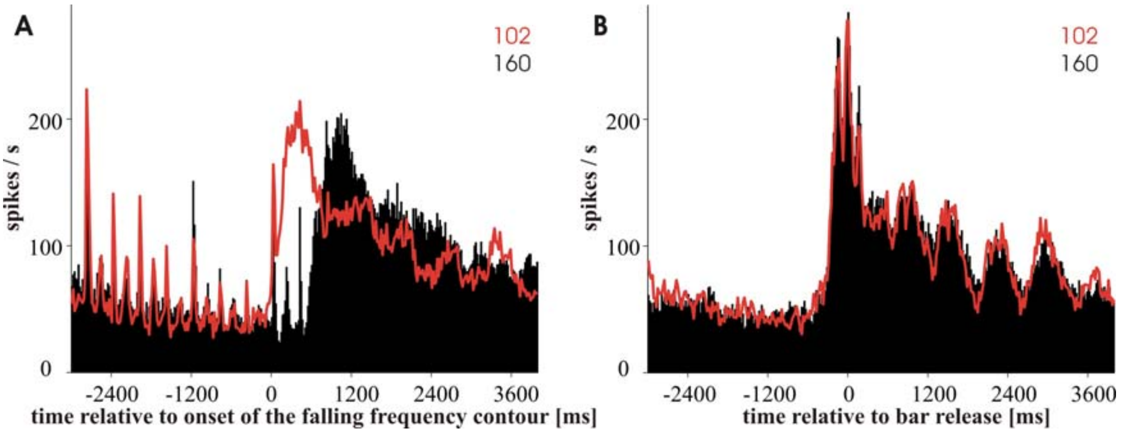

Figure 23: Identification of events which the responses of a multiunit were synchronized to, using trials with different monkey’s reaction time. Red curves show trials in which monkey released the touch bar early (reaction time < 600 ms), black bars show trials in which monkey released the touch bar late (reaction time > 800 ms).

small peaks before bar release, representing firing synchronized to the onset of the tones in the sequence, are not in register. Hence only the three narrow peaks around bar release and the following broader peaks are independent of reaction time and therefore represent firing that is time-locked to bar release.

Firing that was synchronized to the release of the touch bar was found in 36 single units (60 %) and 268 multiunits (85.1 %), similar to the number of acoustically responsive single (37; 61.7 %) and multiunits (271; 86.0 %). Firing started at the earliest 380 ms before the release (median latency -120 ms for multiunits and -100 ms for single units) and lasted, with variable time courses, for a median period of 900 ms in multiunits and 760 ms in single units (Fig. 24). All but two of these units increased their firing during this period, with changes ranging from 60 to 6300 % of the baseline firing (median change 218 %). Release-related multiunit firing was significantly stronger than light-related multiunit firing (median increase 139 %; Wilcoxon-test, $p < 10^{-9}$) and grasping-related multiunit firing (median increase 167 %; $p < 10^{-13}$), when the largest bins in the PETHs were compared. Note that some weak sounds (~ 15 dB SPL) were synchronized with bar release but their measurements exclude them as the sole source of release-related firing.

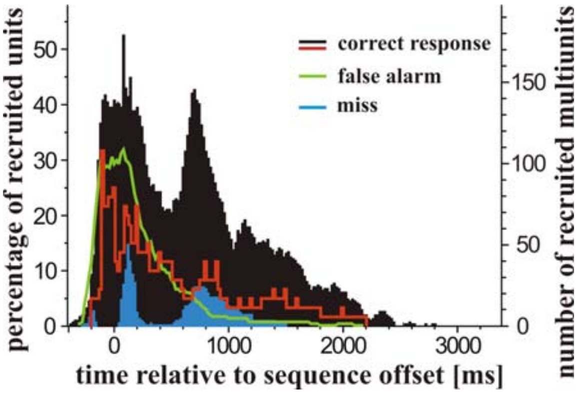

Figure 24: Modulations in neuronal firing before monkeys release the touch bar. Recruitment of units expressed as the percentage or number of units with a significantly modified firing rate in a given bin relative to bar touch. The black bars and the red curve give recruitment of multi and single units for trials with correct responses. Green curve shows recruitment of multiunits in false alarm trials, blue bars correspond miss trials.

By analysis of error trials it was found that in false alarm trial (Fig, 24, green curve), most multiunits started firing before sequence offset. The number of firing multiunits remained high until 140 ms after sequence offset and then decayed monotonically, unlike in correct trials. In miss trials (Fig, 24, blue bars) in which offset of the tone sequence was not coupled with bar release, the first peak in the histogram (at -200 ms) represents responses to last tone in the sequence. However, there were also multiunits that fired after sequence offset. Their firing commenced at the earliest 40 ms after sequence offset, and most multiunits fired 20 ms later. The activation was relatively brief in most multiunits (median duration of 60 ms). As at sequence offset only the cue-light was extinguished this firing may have been evoked by this visual event. In addition to the early firing, there were multiunits that fired during the period 260-1260 ms after sequence offset, most of which ~ 760 ms. The later firing might be related to the release of the touch bar in no-response trials, which occurred, on average, 600 (± 120) ms after sequence offset. Qualitatively similar results were seen in single units although the small number of incorrect trials associated with smaller number of spikes did not permit quantitative analyses.

3.3.3. Relationships between firing related to auditory and non-auditory events

Units in the analyzed sample could exhibit firing that was related to one or more events of the behavioral procedure. This suggests that there may be different types of neurons in auditory cortex namely purely auditory, non-auditory (i.e., neurons that did not respond to pure tones during and outside the behavioral task; no other auditory stimuli were tested), and mixed-type. In the entire sample there were 37 single units with tone responses. A surprisingly small proportion of them (41 %), however, was purely auditory whereas most of them were of mixed type, i.e., they also exhibited firing that was related to the grasp or release of the touch bar. The mixed-type neurons were supplemented

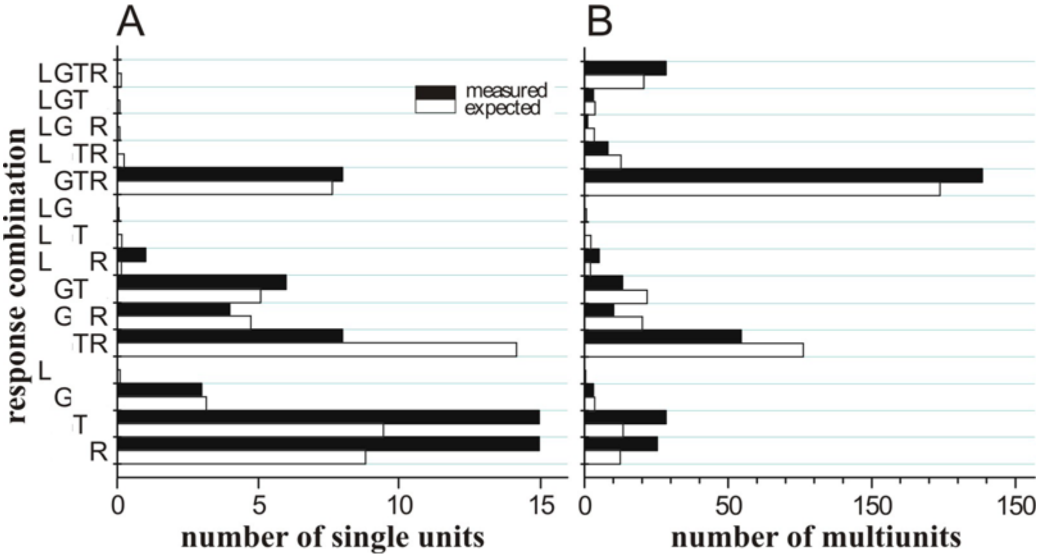

Figure 25: Relationship between different neuronal response properties in single units (A) and multiunits (B). Letters denote firing related to the cue-light (L), grasping (G), tone sequence (T), and bar release (R). Black bars denote the actual numbers that were observed. White bars denote the number of cases that were expected under the assumption that the four properties were statistically independent.

by 23 non-auditory single units. The specific pairings of different response properties are summarized in Fig. 25, A. A Monte Carlo simulation revealed that the pairings were not by chance but systematic. For this simulation, the expected number of units was calculated with different pairings of response properties if pairings were by chance and generated many simulated numbers of units of such chance pairings. Then the average squared difference between the measured and the expected number of pairings was compared to the distribution of the average squared difference between simulated numbers and the expected number with such pairings. This comparison revealed that it was highly unlikely to find the observed pairings in a sample with chance pairings ($p < 0.0001$).

A qualitatively similar though quantitatively different picture emerged when the multiunit data were analyzed (Fig. 25, B). In the multiunit sample, purely auditory sites as well as non-auditory sites were less common than in the single unit sample whereas the number of mixed-type sites was increased.

For the multiunit sample, spectral receptive field properties of units that exhibited non-auditory events related firing were compared with them of purely auditory units (Tab. 1). Receptive field properties were taken from the responses to 40 different pure tones in the passive listening condition and included best frequency as well as the lowest and highest frequency that yielded a response. Pairwise t-tests were applied for each receptive field property to find out whether purely auditory units had response properties different from those of mixed-type or non-auditory units. Except that units without grasping-related firing responded to significantly lower frequencies than units with grasping-related firing, no other significant differences in spectral receptive field properties were found between purely auditory units and units that fired during non-acoustic events.

	light	no light	grasping	no grasping	release	no release
Best frequency [kHz]	11.8	11.1	12.0	11.1	12.0	10.0
Lower frequency border [kHz]	3.9	3.4	4.0	3.4	4.0	2.5
Upper frequency border [kHz]	21.5	21.2	21.4	21.2	21.4	21.0

Table 1: Comparison of tuning curve parameters in multiunits with and without non-auditory firing. Shown are mean values. A significant difference was found for the lower frequency border between units with and without grasping-related firing (pairwise t-tests, $p = 0.0035$). All other differences were not significant with $p > 0.05$.

3.4. Discussion

3.4.1. Category-related firing

Both on the level of individual units and on the population level, phasic responses at short latencies were stronger to falling frequency contours than to rising and flat frequency contours. The following discussion is devoted to the topic how the frequency contour sensitivity of neurons in auditory cortex relates to the category-specificity of other neurons in auditory cortex and to that of neurons in inferotemporal cortex (ITC), lateral prefrontal cortex (PFC), secondary somatosensory cortex SII, ventral premotor cortex PMC, and in the medial temporal lobe.

Are frequency contour sensitive neurons category specific?

Generally, investigators have considered a neuron to be category-selective if the neuronal response pooled over all stimuli belonging to the same category is significantly different from the firing without stimulation and from the responses to stimuli of other categories (e.g., Kreiman et al., 2000). By applying this criterion, 14 category selective multiunits and one category selective single unit were found. All of them responded more strongly to falling frequency contours than to rising and flat frequency contours. The percentage of such units in auditory cortex (24 %) is within the percentage range of category-selective neurons that have been found in other studies (Freedman et al., 2002, 2003: 9 to 26 % in inferotemporal cortex (ITC) and 11 to 32 % in lateral prefrontal cortex (PFC); Sigala and Logothetis, 2002: 34 % in ITC; Vogels (1999): 19 % in ITC; Romo et al., 2002: 17 to 54 % in secondary somatosensory cortex SII; Romo et al., 2004: 44 % in ventral premotor cortex PMC; Kreiman et al., 2000: 72 % in medial temporal lobe of humans). There are several factors that potentially affect the number of category-specific neurons that are found in a study. These include methodological differences and differences in the degree of category-specificity between early and late sensory cortex and prefrontal cortex.

In the present study 20 exemplars from three stimulus classes were tested, namely 8 falling, 8 rising, and 4 flat frequency contours. With this set of stimuli there was found that some units responded more strongly to falling frequency contours than to other contours and that they did so over an 1.5-octave range and for 2 frequency step sizes. The preference for falling frequency contours was not perfect, however. When the individual response magnitudes to all 20 exemplars of frequency contours were compared it could be observed that the 8 strongest responses were not all elicited by falling frequency contours. Rather, among them were always one or more responses that were elicited by rising and, rarely, by flat frequency contours. The reason for this 'contamination' of category-specific units with responses to stimuli of other stimulus classes was that, in addition to frequency contours, neurons in auditory cortex simultaneously encode other features of acoustic signals. One feature is the absolute frequency of the tones of a frequency contour. Therefore the same frequency range for the second tone in the three types of frequency contours was used. Another feature shaping responses of auditory cortex neurons was the size of the frequency interval between the tones of a frequency contour. In correspondence with previous findings on the tone sequence sensitivity of neurons in auditory cortex (Calford and Semple, 1995; Brosch and Schreiner, 1997, 2000; Brosch et al., 1999; Brosch and Scheich, submitted), the responses to the second tone of a pair were generally larger for 1-octave steps than for 0.5-octave steps (this extends previous findings by showing that, at least in actively listening animals, for tone onset asynchronies as long as 400 ms the response to the second tone of a pair is the larger the larger is the frequency separation between sequential tones). Therefore some of the responses to 1-octave upward steps were larger than the responses to 0.5-octave downward steps. The influence of the size of the frequency step could lead to sometimes greater responses to large upward frequency steps than to small downward frequency steps. For this reason, the tests of the putative category specificity of neurons were performed on the same step sizes for rising and falling frequency contours.

If it has been documented, many category-specific neurons described in previous studies also did not appear to be perfectly category specific in the sense that they exclusively responded to stimuli of the same stimulus class. Vogels (1999) noted that the "responses of single neurons in ITC were insufficiently stimulus invariant to accommodate the entire range of variability present in the features of exemplars within the same category". In all 5 examples shown in the articles by Freedman and colleagues (2002, 2003) category-specific neurons in ITC and PFC also responded strongly to some exemplars of other categories. Only in human medial temporal lobe neurons were observed that responded exclusively to stimuli from one category and not to stimuli from other tested categories (Kreiman et al., 2000). It also seems that none of these and other (Sigala and Logothetis, 2002; Romo et al., 2002, 2004) studies has explored possible influences of basic stimulus properties, like the energy or contrast of stimuli, on category-specific responses. Another factor that also has to be taken into account is the number of stimuli from different stimulus classes. The likelihood of misclassifying a neuron as category-specific increases the fewer test stimuli are used. So the number of test stimuli varied considerably between studies (Vogels, 1999: 30; Freedman et al., 2002; 2003: 27; Sigala and Logothetis, 2002: 3 different values in each of 4 feature dimensions; Romo et al., 2002, 2004: > 6; Kreiman et al., 2000: 4-10; present study: 8). In conclusion, these considerations suggest that the category-specificity of the frequency-contour selective units found in the present study is comparable to that described in previous studies.

Influence of the categorization task on the category specificity of neurons

Another factor that could shape the type and degree of category-sensitivity of neuronal responses is the difficulty and the behavioral contingencies of the categorization tasks. Generally, task difficulty can be estimated by assessing the time period required for task acquisition, final performance, and reaction times. Application of these criteria suggests that the categorization of frequency contours was highly difficult and demanding for monkeys. As described in the first part of the present study, about 100,000 trials were required until monkeys performed above chance and they hardly ever scored correctly in more than 75% of the trials. Reaction times were > 700 ms. The categorization tasks used in other studies seem to be more easy for monkeys. Monkeys required about 30,000 trials for the acquisition of a vibrotactile task, in which they had to indicate whether the second of two sequentially presented vibrations was higher or lower in frequency than the first vibration, eventually scored correct in more than 90 % of the trials, and had reaction times < 600 ms (Romo et al., 2002, 2004). Performances in the visual categorization tasks ranged between 85 % in the study of Sigala and Logothetis (2002), and > 90 % in the studies of Freedman and colleagues (2002, 2003), to > 99 % in the study of Vogels (1999). Reaction times were < 200 ms (Vogels, 1999) and ~ 300 ms (Freedman et al., 2003). The number of trials required to reach criterion was not given in any of these studies.

The exclusive preference for one stimulus class in the auditory cortex described in the present study is unlikely due to differences between early sensory cortex and higher level cortical areas. In the auditory cortex of non-behaving monkeys (Steinschneider et al., 1982, 1995) or anesthetized cats (Eggermont, 1995) it was found that categorically perceived voiced and unvoiced stop-consonant/vowel syllables were reflected by different temporal coding mechanisms, such that the stimuli from both stimulus classes elicited transient response bursts that were time-locked to consonant release only or also to voicing onset. The type of category-specificity of neurons described in the present study is more likely due to the behavioral contingencies of the categorization task. This could be the main reason why, in contrast to all previous studies, all frequency contour sensitive neurons in auditory cortex responded more strongly to stimuli of the same class, namely to falling frequency contours. In the present experiments monkeys were required to release a touch bar immediately after the occurrence of this type of frequency contour. Thus only one stimulus class prompted a motor act and was associated with reward. Like in the present study, Freedman and colleagues (2002, 2003) also used a go/nogo behavioral paradigm in which monkeys watched a sequence of two pictures and were required to release a touch bar when the pictures belonged to the same stimulus class and to maintain contact otherwise. In contrast to the present study, however, both classes of stimuli could be go-stimuli which means that both classes of stimuli had the same behavioral relevance. Two-alternative forced choice procedures were also used in the experiments performed by Vogels (1999), Sigala and Logothestis (2002), and Romo and colleagues (Romo et al., 2002, 2004), i.e., all stimuli were associated with motor acts (saccades in different directions or pressing different push buttons). It could be supposed that rising frequency contours would elicit stronger responses than falling frequency contours if monkeys signaled the former contour type.

3.4.2. Non-auditory event related firing

Many neurons in auditory cortex fired also during non-acoustic events of the behavioral procedure (Brosch et al., 2005). Because most of the firing was transient and time-locked to these events and was not due to artifacts, the conclusions are: (1) The firing after turning on the cue-light was evoked by this light. (2) The firing before contact with the touch bar was related to the preparation or execution of movements. (3) The firing before the bar release was not related to the recognition of falling frequency contour in the tone sequence.

These results extend previous findings that auditory cortex responses to acoustic stimuli are modulated during the performance of an auditory task (Miller et al., 1972; Benson and Hienz, 1978; Fritz et al., 2003, 2005). They corroborate that auditory cortex neurons respond to somatosensory stimulation in awake non-performing (Schroeder et al., 2001) and anesthetized monkeys (Fu et al., 2003) and to visual stimuli in awake non-performing (Schroeder and Foxe, 2002) and in experimentally cross-wired animals (Sur et al., 1990). These findings are in line with recent demonstrations of anatomical connections between early auditory and visual cortical areas (Falchier et al., 2002; Rockland and Ojima, 2003) and complement reports of acoustic responses in visual cortical areas 17-19 (Morrell, 1972; Fishman and Michael, 1973; Bronchti et al., 2002).

The sustained changes of firing during the bar holding period may reflect stimulus expectation or preparation and adjustments for the processing of upcoming acoustic stimuli. The transient activation before and synchronized to bar release may, in part, reflect late responses to falling frequency contour that are modified in latency and magnitude by the monkeys' decision to release the bar. By contrast, modulatory inputs are unlikely to account for most of the firing in auditory cortex that was synchronized with non-acoustic events of the behavioral task, like the onset of the cue-light and the grasping and release of the touch bar. First, during their occurrence there were no responses to acoustic stimuli that could have been modulated by non-auditory input. Second, many transient responses to non-acoustic events occurred at short latencies, comparable to those in early visual (Raiguel et al., 1989) and somatosensory cortical areas (Romo et al., 1996).

The responses to non-acoustic stimuli and the firing during arm movements suggest that auditory cortex can represent non-acoustic events, in addition to sound. The interpretation of the extensive multimodal co-representation in the auditory cortex of highly trained monkeys is that the categorization of acoustic stimuli was intimately associated with visual stimuli and hand movements. Similar audio-motor integration at behavioral and cortical levels was observed in professional musicians (Münte et al., 2002). There is also some evidence that blind subjects benefit from a somatosensory representation in visual cortex for Braille reading (Cohen et al., 1997). It is conceivable that the co-representation of non-acoustic events in the auditory cortex of the monkeys has emerged during the long training period they have spent on the acquisition of the task. The representation of non-auditory sensory modalities and movements in auditory cortex could accelerate and improve the subject's performance in highly demanding auditory tasks (Bangert et al., 2001).

4. General Discussion

To study neuronal mechanisms underlying the categorical perception of frequency contours, single and multiunit firing were recorded from the auditory cortex of highly trained monkeys while they performed a go/nogo positive reinforcement behavioral paradigm in which they signaled the occurrence of falling frequency contours in tone sequences. With respect to category learning theories, the frequency-direction task qualifies as a rule-based type (Ashby and Shawn, 2001) but not only because a simple abstract rule of distinguishing upward from downward steps leads to the categories. The monkeys could also use the procedural rule that the identification of the first frequency change in this paradigm predicted all subsequent events. Furthermore, because transient responses identified no frequency band in which categorical responses were preferred, it is likely that an exemplar model rather than a prototype model accounts for the accumulation of category information. Chiefly from deficits in patients rule-based category learning is attributed usually to frontal-striatal and anterior cingulate cortex circuits (Ashby et al., 1998; Elliott et al., 1999). The present study suggests that these circuits interact with auditory cortex which in a highly trained state reflects decision making and not only contributes sensory information.

The selective increase of neuronal responses to falling frequency contours seems to be due to the behavioral salience of this feature in the tone sequences, namely that this feature prompts a motor act (bar release), which can result in reward. This interpretation is in line with and extends previous observations that specific pairings of sensory stimuli with motor acts, with positive or negative reinforcement, or with electrical stimulation of various deep brain structures can result in selective changes of the responses to this and adjacent stimuli (Ma and Suga, 2003; Weinberger, 2004; Ohl and Scheich, 2005). In many cases, it was found that such pairings selectively enhance neuronal responses to the frequency of a positively conditioned tone and relatively non-selectively reduces responses to all other frequencies, including negatively conditioned frequencies (Ma and Suga, 2003; Weinberger, 2004; Blake et al, 2002, Fritz et al, 2003). Other studies demonstrated that response enhancement of specific tones reflected more global effects in which the sensitivity of the local slopes of frequency tuning curves around the trained frequencies was increased (Ohl and Scheich, 1996, 2005). Aside from frequency-selective changes of neuronal responses, changes in neuronal responses also have been found for other acoustic stimuli, such as periodically modulated tones (Beitel et al, 2003; Ohl at al., 2001), click trains (Bao et al, 2004), and intensity variations (Polley et al., 2004). These modulations are not limited to the auditory cortex but include subcortical structures of the auditory system (Gonzales-Lima and Scheich, 1986; Ma and Suga, 2003; Weinberger, 2004) and have also been observed with visual stimuli in higher visual areas (lateral intraparietal area, LIP; Gottlieb et al, 1989) and PFC (Watanabe, 1990, 1992; Assad et al, 1998). As the remapping of frequency contours to motor acts was present from the beginning of the responses to the tones with the latency range of 20 to 40 ms, it is possible that some form of increased sensitivity for falling frequency contours is already present at subcortical stages of the auditory system. Results of the present study extend these findings by showing that stimulus-response mapping by enhanced responses does not only occur for individual stimuli but can also occur for an entire class of stimuli, like for falling frequency contours.

The mechanisms generating enhanced responses to falling frequency contours relative to other contours could arise from modifications of sequential two-tone interactions, such that the frequency range that induces poststimulatory response inhibition changes relative to the poststimulatory response inhibition seen outside the behavioral task or in naive animals (Brosch et al., 1998). Stronger responses to falling frequency contours could result from weaker inhibition induced by tones with frequencies above central frequency or by stronger inhibition induced by tones with frequencies below central frequency. Although, in principle, the same effects could result from changes of poststimulatory response facilitation there is less likely because enhanced responses to falling frequency contours occurred from the very beginning of the tone responses. In monkey auditory cortex poststimulatory response facilitation, by contrast, affects later parts of the neuronal responses after the initial phasic response (Brosch et al., 1999).

The relative enhancement of responses to falling frequency contours seems to be a form of long-term response plasticity of neurons in auditory cortex. This is in contrast to rapid response plasticity which is characterized as modulation of response properties of neurons that is driven by attentive focus on a salient acoustic feature and occurs within minutes or shorter, long-term response plasticity results in permanent transformations of response properties of neurons compared to such properties before experimental treatment or to those in naive animals. Analysis of error trials shows that also by incorrect behavioral response responses to falling frequency contours are stronger than these to rising frequency contour. However, in false alarm trials responses to rising frequency contour are significantly stronger than responses to rising frequency contour in correct trials (Fig. 16). One possible explanation is that this response enhancement to rising frequency contour causes the incorrect motor reaction. On the other hand, the strong transient response may not be the sole trigger of the motor reaction because the responses to the initial tone are for example stronger than responses to the other tones in the sequence but don’t lead to the motor reaction. The additional mechanism seems to be provided from the slow modulations of neuronal firing during the task. So the slow increase of firing rate after monkeys grasp the touch bar could be such kind of mechanism which leads to the additional activation in the moment in which monkeys must categorize presented stimuli.

Although the training on frequency contours likely has induced permanent modifications of the responses of the neurons in auditory cortex of the experimental monkeys, generally the preference for falling frequency contours seems to be task-related. Analyze of sequential two-tone interactions in naive awake (Brosch and Scheich, submitted) and anesthetized monkeys (Brosch et al, 1999; Brosch and Scheich, 2003) showed that responses of neurons to a downward frequency step were indistinguishable from responses to an upward frequency step of the same size, which is in contrast to the response asymmetry found during task performance in the present experiments. Interestingly performance of the auditory discrimination task also resulted in an increase of the time-constant of sequential two-tone interactions, compared to the non-performing state. These task-related response modulations reflect the highly adaptive character of auditory cortex.

In addition, some of the firing seen in the present experiments may be prone to non-sensory modulations. Both grasping-related and release-related firing in auditory cortex appeared to require the monkey to be engaged in an auditory task because these activations disappeared when the monkey performed a visual instead of the standard auditory task. This may also apply to cue-light related firing in auditory cortex, which was hardly ever observed when the monkeys subsequently did not touch the bar. This implies that, in many neurons, cue-light related firing was evoked only under specific conditions, as when the monkey attended to visual stimuli, associated the cue-light with auditory processing, or when the visual stimulus fell into the visual receptive field of a cell in auditory cortex.

An explanation for these non-acoustical modulations could be provided by the functional system theory (Alexandrov and Jarvilehto, 1993; Alexandrov et al., 1997). According to this theory, all the elements of the organism are organized in systems that are neither sensory nor motor, but functional, and the elements of these systems are defined in terms of how they enable the achievement of useful results of behavior i.e. taking food, avoiding harmful effects or breathing. The result is a new relation between the organism and environment, an achievement of a concrete adaptive change in the interaction between the organism and environment. From this point of view, the monkey’s task performance is mediated by one of the functional systems and aimed to take water. Non-acoustical modulations in the auditory cortex are due to a leading role of the acoustical stimulation for the useful result of behavior. The cue-light indicated that the monkeys could initiate a tone sequence. By grasping and holding the touch bar the monkeys signaled their readiness to listen to acoustic stimuli and started the tone sequence. By withdrawing their hand from the touch bar the monkeys expressed the result of their auditory processing, what in successful trials leaded to reward.

5. References

Ahissar E, Vaadia E, Ahissar M, Bergman H, Arieli A, Abeles M (1992) Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. Sciense 257: 1412-1415.

Aizenstein HJ, MacDonald AW, Stenger VA, Nebes RD, Larson JK, Ursu S, Carter CS (2000) Complementary category learning systems identified using event-related functional MRI. J Cogn Neurosci. 12(6): 977-987.

Alexandrov YuI, Grechenko TN, Gavrilov VV, Gorkin AG, Shevchenko DG, Grinchenko YuV, Aleksandrov IO, Maksimova NE, Bezdenezhnych BN, Bodunov MV (1997) Formation and realization of individual experience. Neurosci Behav Physiol 27(4):441-54.

Alexandrov YuI, Jarvilehto T (1993) Activity versus reactivity in psychology and neurophysiology. Ecological Psychol 5(1), 85-103.

Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM (1998) A neuropsychological theory of multiple systems in category learning. Psychol Rev 105(3): 442-481.

Ashby FG, Gott RE (1988) Decision rules in the perception and categorization of multidimensional stimuli. J Exp Psychol Learn Mem Cogn 14(1): 33-53.

Ashby FG, Shawn WE (2001) The neurobiology of human category learning. Trends Cogn Sci 5 (5): 204-210.

Assad WF, Rainer G, Miller EK (1998) Neural activity in the primate prefrontal cortex during associative learning. Neuron 21:1399–1407.

Bangert M, Haeusler U, Altenmuller E (2001) On practice: how the brain connects piano keys and piano sounds. Ann N Y Acad Sci 930: 425-428.

Bartlett EL, Wang X (2005) Long-lasting modulation by stimulus context in primate auditory cortex. J Neurophysiol 94: 83-104.

Beaton R, Miller JM (1975) Single cell activity in the auditory cortex of the unanesthetized, behaving monkey: correlation with stimulus controlled behavior. Brain Res 100: 543-562.

Beitel RE, Schreiner CE, Cheung SW, Wang X, Merzenich MM (2003) Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. Proc Natl Acad Sci USA 100: 11070-11075.

Benson DA, Hienz RD (1978) Single-unit activity in the auditory cortex of monkeys selectively attending left vs. right ear stimuli. Brain Res 159: 307-320.

Blake DT, Strata F, Churchland AK, Merzenich M (2002) Neural correlates of instrumental learning in primary auditory cortex. PNAS 99(15): 10114-10119.

Braaten RF, Hulse SH (1993) Perceptual organization of auditory temporal patterns in European starlings (sturnus vulgaris). Percept Psychophys, 54(5): 567-578.

Braaten RF, Hulse SH, Page SH (1990) Discrimination and classification of rising and nonrising pitch patterns by the European starling (sturnus vulgaris). Anim Learn Behav, 18: 352-364.

Bregman AS (1990) Auditory scene analysis. The perceptual organization of sound. MIT Press, Cambridge, London.

Bronchti G, Heil P, Sadka R, Hess A, Scheich H, Wollberg Z (2002) Auditory activation of "visual" cortical areas in the blind mole rat (Spalax ehrenbergi). Eur J Neurosci 16: 311-329.

Brosch M, Scheich H (submitted) Tone-sequence analysis in the auditory cortex of awake Macaque monkeys.

Brosch M, Selezneva E, Bucks C, Scheich H (2004) Macaque monkeys discriminate pitch relationships. *Cognition.* 91(3):259-72.

Brosch M, Selezneva E, Scheich H (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25(29):6797-806.

Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. J Neurophysiol 77(2): 923-943.

Brosch M, Schreiner CE (2000) Sequence sensitivity of neurons in cat primary auditory cortex. Cereb Cortex 10(12): 1155-1167.

Brosch M, Schulz A, Scheich H (1998) Neuronal mechanisms of auditory backward recognition masking in macaque auditory cortex. Neuroreport 9(11):2551-2555.

Brosch M, Schulz A, Scheich H (1999) Processing of sound sequences in macaque auditory cortex: response enhancement. J Neurophysiol 82: 1542-1559.

Brown RG, Marsden CD (1988) Internal versus external cues and the control of attention in Parkinson’s disease. Brain 111: 323-345.

Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. J Neurophysiol 73(5): 1876-1891.

Chang HW, Trehub SE (1977) Auditory processing of relational information by young infants. J Exp Child Psychol, 24: 324-331.

Cohen LG, Celnik P, Pascual-Leone A, Corwell B, Falz L, Dambrosia J, Honda M, Sadato N, Gerloff C, Catala MD, Hallett M (1997) Functional relevance of cross-modal plasticity in blind humans. Nature 389: 180-183.

Cools AR, van den Bercken JH, Horstink MW, van Spaendonck KP, Berger HJ (1984) Cognitive and motor shifting aptitude disorder in Parkinson’s disease. J Neurol Neurosurg. Psychiatry 47(5): 443-453.

Cynx J (1995) Similarities in absolute and relative pitch perception in songbirds (starling and zebra finch) and a nonsongbird (pigeon). J Comp Psychol, 109: 261-267.

D'Amato MR (1988) A search for tonal pattern perception in cebus monkey: Why monkeys can't hum a tune. Music Percept, 5: 452-480.

D'Amato MR, Colombo M (1985) Auditory matching-to-sample in monkeys (Cebus apella). Anim Learn Behav, 13: 375-382.

D'Amato MR, Salmon DP (1984) Processing and retention of complex auditory stimuli in monkeys (Cebus apella). Canadian J Psychol, 38: 237-255.

Davis H, Albert M (1986) Numerical discrimination by rats using sequential auditory stimuli. Anim Learn Behav, 14: 57-59.

Diamond DM, Weinberger NM (1986) Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. Brain Res 372: 357-360.

Diamond DM, Weinberger NM (1989) Role of context in the expression of learning-induced plasticity of single neurons in auditory cortex. Behav Neurosci 103: 471-494.

Dowling WJ (1978) Scale and contour: two components of a theory of memory for melodies. Psychol Rev, 85: 341-354.

Durif C, Jouffrais C, Rouiller EM (2003) Single-unit responses in the auditory cortex of monkeys performing a conditional acousticomotor task. Exp Brain Res 153: 614-627.

Eggermont JJ (1995) Representation of a voice onset time continuum in primary auditory cortex of the cat. J Acoust Soc Am 98: 911-920.

Eliades SJ, Wang X (2003) Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol 89: 2194-2207.

Eliades SJ, Wang X (2005) Dynamics of auditory-vocal interaction in monkey auditory cortex. Cereb Cortex15(10):1510-1523.

Elliott R, Rees G, Dolan RJ (1999) Ventromedial prefrontal cortex mediates guessing. Neuropsychologia 37(4):403-411.

Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22: 5749-5759.

Filoteo JV, Maddox WT, Davis JD (2001) A possible role of the striatum in linear and nonlinear category learning: evidence from patients with Huntington’s disease. Behav Neurosci 115(4): 786-798.

Finney, EM Fine I, Dobkins KR (2001) Visual stimuli activate auditory cortex in the deaf. Nat Neurosci 4: 1171-1173.

Fishman MC, Michael P (1973) Integration of auditory information in the cat's visual cortex. Vision Res 13: 1415-1419.

Fitch RH, Brown CP, O'Connor K, Tallal P (1993) Functional lateralization for auditory temporal processing in male and female rats. Behav Neurosci, 107: 844-850.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2002) Visual categorisation and the primate prefrontal cortex: neurophysiology and behavior. J Neurophysiol 88: 929-941.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23(12):5235-5246.

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci 6(11): 1216-1223.

Fritz J, Elhilali M, Shamma S (2005) Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. Hearing Res 206(1-2): 159-176.

Fu KM, Johnson TA, Shah AS, Arnold I, Smiley J, Hackett TA, Garraghty PE, Schroeder CE (2003) Auditory cortical neurons respond to somatosencory stimulation. J Neurosci 23: 7510-7515.

Fukushi T, Sawaguchi T (2005) Neuronal representation of response category and motor parameters in monkey prefrontal cortex. Exp Brain Res, 164(4): 472-483.

Gonzalez-Lima F, Scheich H (1986) Classical conditioning of tone-signaled bradycardia modifies 2-deoxyglucose uptake patterns in cortex, thalamus, habenula, caudate-putamen and hippocampal formation. Brain Res. 363: 239-256.

Gottlieb Y, Vaada E, Abeles M (1989) Single unit activity in the auditory cortex of a monkey performing a short term memory task. Exp Brain Res 74: 139-148.

Griffiths TD (2003) Functional imaging of pitch analysis. Ann NY Acad Sci 999: 40-49.

Hampson M, Olson IR, Leung HC, Skudlarski P, Gore JC (2004) Changes in functional connectivity of human MT/V5 with visual motion input. Neuroreport 15(8):1315-1319.

Harnad S (1987) Categorical perception. The groundwork of cognition. Cambrodge University Press, Cambrodge.

Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. Cognition, 78: B53-64.

Hocherman S, Itzhaki A, Gilat E (1981) The response of single units in the auditory cortex of rhesus monkeys to predicted and to unpredicted sound stimuli. Brain Res 230: 65-86.

Homa D, Sterling S, Trepel L (1981) Limitations of exemplar-based generalization and the abstraction of categorical information. J Exp Psychol Hum Learn Mem 7:418-439.

Hubel DH, Henson CO, Rupert A, Galambos R (1959) Attention units in the auditory cortex. Science 129: 1279-1280.

Isomura Y, Ito Y, Akazawa T, Nambu A, Takada M (2003) Neural coding of ‘attention for action’ and ‘response selection’ in primate anterior cingulated cortex. J Neurosci 23(22): 8002-8012.

Izumi A (2001) Relative pitch perception in Japanese monkeys (Macaca fuscata). J Comp Psychol, 115(2): 127-131.

Knowlton BJ, Ramus SJ, Squire LR (1992) Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. Psychol Sci 3: 172-179.

Knowlton BJ, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. Learn Mem 1(2): 106-120.

Knowlton BJ, Squire LR, Paulsen JS, Swerdlow NR, Swenson M, Butters N (1996) Dissociations within nondeclarative memory in Huntington’s disease. Neuropsychology 10(4): 538-548.

Kojima S (1985) Auditory short-term memory in the Japanese monkey. Int J Neurosci, 25: 255-262.

Kolodny JA (1994) Memory processes in classification learning: an investigation of amnestic performance in categorization of dot patterns and artistic styles. Psychol Sci 5: 164-169.

Kreiman G, Koch C, Fried I (2000) Category-specific visual responses of single neurons in the human medial temporal lobe. Nat Neurosci 3(9): 946-953.

Lee TS, Yang CF, Romero RD, Mumford D (2002) Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. Nat Neurosci 5: 589-597.

Lerdahl F, Jackendoff R (1983) A Generative Theory of Tonal Music. MIT Press, Cambridge.

Ma X, Suga N (2003) Augmentation of plasticity of the central auditory system by the basal forebrain and/or somatosensory cortex. J Neurophysiol 89: 90-103.

MacDougall-Shackleton SA, Hulse SH (1996) Concurrent absolute and relative pitch processing by European starlings (Sturnus vulgaris). J Comp Psychol 110: 139-146.

Maddox WT, Filoteo JV (2001) Striatal contribution to category learning: quantitative modeling of simple linear and complex non-linear rule learning in patients with Parkinson’s disease. J Int Neuropsychol Soc 7(6):710-727.

Merchant H, Zainos A, Hernandez A, Salinas E, Romo R (1997) Functional properties of primate putamen neurons during the categorization of tactile stimuli. J Neurophysiol 77(3): 1132-1154.

Meulemans T, Peigneux P, van der Linden M (1998) Preserved artificial grammar learning in Parkinson’s disease. Brain Cognit 37: 109-112.

Miller JM, Dobie RA, Pfingst BE, Hienz RD (1980) Electrophysiologic studies of the auditory cortex in the awake monkey. Am J Otolaryngol 1(2): 119-130.

Miller JM, Sutton D, Pfingst B, Ryan B, Ryan A, Beaton R, Gourevitch G (1972) Single cell activity in the auditory cortex of Rhesus monkeys: behavioral dependency. Science 177: 449-451.

Morrell F (1972) Visual system's view of acoustic space. Nature 238: 44-46.

Mueller-Preuss P, Ploog D (1981) Inhibition of auditory cortical neurons during phonation. Brain Res 215: 61-76.

Muente TF, Altenmueller E, Jaencke L (2002) The musician's brain as a model of neuroplasticity. Nat Rev Neurosci 3: 473-478

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) ‘Primitive intelligence’ in the auditory cortex. Trends neurosci 24(5): 283-288.

Ohl FW, Scheich H (1996) Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field AI) of the alert Mongolian gerbil. Eur J Neurosci. 8: 1001-1017.

Ohl FW, Scheich H (2005) Learning-induced plasticity in animal and human auditory cortex. Curr Opin Neurobiol 15(4): 470-477.

Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. Nat 412:733-736.

Paavilainen P, Jaramillo M, Näätänen R, Winkler I (1999) Neuronal populations in the human brain extracting invariant relationships from acoustic variance. Neurosci Lett. 265(3): 179-182.

Page SH, Hulse SH, Cynx J (1989) Relative pitch perception in the European starling (sturnus vulgaris): Further evidence for an elusive phenomenon. J Exp Psychol Anim Behav Process, 15: 137-146.

Pekkola J, Ojanen V, Autti T, Jaaskelainen IP, Mottonen R, Tarkiainen A, Sams M (2005) Primary auditory cortex activation by visual speech: an fMRI study at 3 T. Neuroreport 16(2): 125-128.

Pfingst BE, O’Connor TA, Miller JM (1977) Response plasticity of neurons in auditory cortex of the rhesus monkey. Exp Brain Res 29: 393-404.

Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, Poeppel D, McGinnis M, Roberts T (2000) Auditory cortex accesses phonological categories: an MEG mismatch study. J Cogn Neurosci 12(6): 1038-1055.

Posner MI, Keele SW (1968) On the genesis of abstract ideas. J Exp Psychol 77:353-363.

Rao SM, Bobholz JA, Hammeke TA, Rosen AC, Woodley SJ, Cunnungham JM, Cox RW, Stein EA, Binder JR(1997) Functional MRI evidence for subcortical participation in conceptual reasoning skills. Neuroreport 8(8): 1987-1993.

Raiguel SE, Lagae L, Gulyas B, Orban GA (1989) Response latencies of visual cells in macaque areas V1, V2 and V5. Brain Res 493: 155-159.

Reber PJ, Stark CE, Squire LR (1998) Contrasting cortical activity associated with category memory and recognition memory. Learn Mem 5(6): 420-428.

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13(1):87-103.

Rockland KS, Ojima H (2003) Multisensory convergence in calcerine visual areas in macaque monkey. Int J Psychophysiol 50: 19-26.

Robinson AL, Heaton RK, Lehman RA, Stilson DW (1980) The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. J Consult Clin Psychol 48(5):605-614.

Rolls ET, Judge SJ, Sanghera MK (1977) Activity of neurons in the inferotamporal cortex of the alert monkey. Brain Res 130: 229-238.

Romo R, Hernandez A, Zainos A, Lemus L, Brody CD. (2002) Neuronal correlates of decision-making in secondary somatosensory cortex. Nat Neurosci 5: 1217-1125

Romo R, Hernandez A, Zainos A. (2004) Neuronal correlates of a perceptual decision in ventral premotor cortex. Neuron 41:165-173.

Romo R, Merchant H, Ruiz S, Crespo P, Zainos A (1995) Neuronal activity of primate putamen during categorical perception of somaesthetic stimuli. Neuroreport 6(7):1013-1017.

Romo R, Merchant H, Zainos A, Hernandez A (1996) Categorization of somaesthetic stimuli: sensorimotor performance and neuronal activity in primary somatic sensory cortex of awake monkeys. Neuroreport 17: 1273-1279.

Romo R, Merchant H, Zainos A, Hernandez A (1997) Categorical perception of somesthetic stimuli: psychophysical measurements correlated with neuronal events in primate medial premotor cortex. Cereb Cortex 7(4): 317-326.

Romo R, Ruiz S, Crespo P, Zainos A, Merchant H (1993) Representation of tactile signals in primate supplementary motor area. J Neurophysiol 70(6): 2690-2694.

Sakurai Y (1994) Involvement of auditory cortical and hippocampal neurons in the auditory working memory and reference memory in the rat. J Neurosci, 14: 2606-2623.

Salinas E, Romo R (1998) Conversion of sensory signals into motor commands in primary motor cortex. J Neurosci 18(1): 499-511.

Schmidt EM (1984) Computer separation of multi-unit neuroelectric data: a review. J Neurosci Methods 12:95-111.

Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC (2001) Somatocensory input to auditory association cortex in the macaque monkey. J Neurophysiol 85: 1322-1327.

Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res Cogn Brain Res 14: 187-198.

Seger CA, Cincotta CM (2002) Striatal activity in concept learning. Cogn Affect Behav Neurosci 2:149-161.

Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. Nat 415: 318-320.

Sinnott JM, Kreiter NA (1991) Differential sensitivity to vowel continua in Old World monkeys (Macaca) and humans. J Acoust Soc Am, 89: 2421-2429.

Stein BE, Meredith MA (1993) The Merging of the Senses. Boston: MIT.

Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. (1994) Speech-evoked activity in primary auditory cortex: effects of voice onset time. Electroencephalogr Clin Neurophysiol 92(1): 30-43.

Steinschneider M, Arezzo J, Vaughan HG Jr. (1982) Speech evoked activity in the auditory radiations and cortex of the awake monkey. Brain Res 252: 353-356.

Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. (1995) Physiologic correlates of the voice onset time boundary in primary auditory cortex (A1) of the awake monkey: temporal response patterns. Brain Lang. 48(3): 326-340.

Suga N, Ma X (2003) Multiparametric corticofugal modulation and plasticity in the auditory system. Nat Rev Neurosci 4: 783-794.

Sur M, Pallas SL, Roe AW (1990) Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex. Trends Neurosci 13: 227-233.

Trehub SE, Bull D, Thorpe LA (1984) Infants' perception of melodies: The role of melodic contour. Child Development, 55: 821-830.

Vaadia E, Gottlieb Y, Abeles M (1982) Single-unit activity related to sensorimotor association in auditory cortex of a monkey. J Neurophysiol 48: 1201-1213.

Vogels R (1999) Effect of image scrambling on inferior temporal cortical responses. Neuroreport 10(9): 1811-1816.

Wetzel W, Ohl FW, Wagner T, Scheich H (1998) Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. Neurosci Lett 252(2):115-8.

Wetzel W, Wagner T, Ohl FW, Scheich H (1998) Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbels. Behav Brain Res 91(1-2): 29-39.

Yoshioka T, Dow BM, Vautin RG (1996) Neuronal mechanisms of color categorization in areas V1, V2 and V4 of macaque monkey visual cortex. Behav Brain Res 76(1-2): 51-70.

Watanabe M. (1990) Prefrontal unit activity during associative learning in the monkey. Exp Brain Res. 80: 296-309.

Watanabe M. (1992) Frontal units of the monkey coding the associative significance of visual and auditory stimuli. Exp Brain Res. 89: 233-247.

Weinberger NM (2004) Specific long-term memory traces in primary auditory cortex. Nat Rev Neurosci 5: 279-290.

Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM (2003) Eye position affects activity in primary auditory cortex of primates. Curr Biol 13: 554-562.

Witte RS, Kipke DR (2005) Enhanced contrast sensitivity in auditory cortex as cats learn to discriminate sound frequencies. Cogn Brain Res 23: 171-184.

Wright AA, Shyan MR, Jitsumori M (1990) Auditory same/difference concept learning in monkeys. Anim Learn Behav, 18: 287-294.

Wright AA, Rivera JJ, Hulse SH, Shyan MR, Neiworth JJ (2000) Music perception and octave generalization in rhesus monkeys. J Exp Psychol General, 129: 291-307.

Wright AA (1991) Concept learning by monkeys and pigeons. In Memory Mechanisms: A Tribute to G. V. Goddard; Abraham WC, Corballis M, White KG (ed.), pp. 247-273, Lawrence Earlbaum Associates, Hillsdale, NJ.

Zaehle T, Wustenberg T, Meyer M, Jancke L (2004) Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. Eur J Neurosci, 20(9):2447-56.

Appendices

I. Zusammenfassung

Kategorisierung ist der Prozess, mit dem Konzepte und Objekte wahrgenommen und verstanden werden. Dieser Vorgang ist fundamental beim Entscheidungsprozess und in allen Arten von Interaktionen mit der Umwelt. In der vorliegenden Arbeit wurde untersucht, wie akustische Kategorien wie melodische Konturen (fallend, eben und steigend) im auditorischen Kortex repräsentiert sind.

Zwei *Macaca fascicularis* wurden trainiert, melodische Konturen in Tonsequenzen zu diskriminieren. Die Tonsequenzen bestanden aus zwei oder drei Serien von Tönen mit verschiedener Frequenz. Innerhalb einer Serie hatten alle Töne dieselbe Frequenz. Die zweite Tonserie hatte entweder eine höhere oder eine tiefere Frequenz als die erste Tonserie. Im Falle einer höheren Frequenz wurde die zweite Tonserie von einer dritten Tonserie tieferer Frequenz gefolgt. Die Affen sollten nach dem Lichtsignal einen Berührungsschalter anfassen und damit eine Tonsequenz starten. Die weitere Aufgabe der Affen bestand darin, den Berührungsschalter loszulassen, sobald ein Ton mit absteigender Frequenz auftauchte. Sie durften den Berührungsschalter nicht loslassen, wenn der folgende Ton dieselbe oder eine höhere Frequenz als der vorhergehende Ton hatte. Nach ca. 100000 Trials konnten beide Affen melodische Konturen auf kategorielle Weise (unabhängig von der Größe des Frequenzintervals und von der Absolutfrequenz der Töne) diskriminieren.

Nach Abschluß des Verhaltenstrainings wurden Mikroelektrodenableitungen im primären auditorischen Kortex und im caudomedialen auditorischen Kortexfeld dieser Affen durchgeführt, während diese weiterhin die akustische Diskriminationsaufgabe durchführten. Es wurde festgestellt, daß neuronale Antworten im Hörcortex dieser hoch trainierten Affen sowohl akustische als auch Reize aus anderen Modalitäten repräsentieren.

Neuronen antworteten phasisch mit einer Latenz von ~ 30 ms auf einzelne, hintereinander folgende Töne in den Tonsequenzen. Die Antworten wurden schwächer, wenn der vorherige Ton dieselbe Frequenz hatte, und stärker, wenn der vorherige Ton höher oder tiefer war. Dieses war abhängig von dem Frequenzintervall zwischen den Tönen (Antworten wurde umso stärker, je unterschiedlicher die Frequenzen der Töne waren) und auch von der Richtung

der Frequenzänderung (vorheriger Ton tiefere Frequenz, rief stärkere Antwort hervor als vorheriger Ton höhere Frequenz). Die Präferenz für fallende Konturen wurde sowohl auf der Populationsebene (Populationsantwort auf abfallende Konturen war ~15 % stärker als Populationsantwort auf aufsteigende Konturen und ~25% stärker als Populationsantwort auf ebene Konturen) als auch auf dem Niveau der einzelnen Zellen (24% der Neuronen antworteten signifikant stärker auf eine fallende Kontur und kein Neuron antwortete signifikant stärker auf eine steigende oder eine ebene Kontur) festgestellt. Dies war unabhängig von den Frequenz der Töne und tauchte sowohl bei richtigen als auch bei falschen Entscheidungen der Affen auf.

Neuronen feuerten aber nicht nur bei der Präsentation der Töne in der Sequenz, sondern auch zu anderen Zeitpunkten der Diskriminationsaufgabe. ~15% der Neuronen antworteten mit einer Latenz von ~100 ms auf das Lichtsignal bevor der Affe den Berührungsschalter anfaßte. Diese Lichtantworten wurde nicht gefunden, wenn der Affe keinen Kontakt zu dem Schalter aufbaute, d.h. keine Bereitschaft zu der akustischen Aufgabe zeigte. ~60% der Neuronen antworteten phasisch (mit einer Latenz von ~ 100 ms) oder tonisch zu dem Zeitpunkt, an dem der Affe den Schalter berührte. Es wurden zwei Typen von tonischer Aktivierung gefunden: 12 % der Neuronen feuerten schwächer, mindestens 140 ms nachdem der Affe den Berührungschalter anfaßte und 13% der Neuronen feuerten stärker, mindestens 1 s nach der Berührung des Schalters. ~ 60% der Neuronen feuerten stärker zu dem Zeitpunkt, an dem der Affe den Kontakt mit dem Schalter unterbrach. Diese Aktivierung startete ~120 ms bevor der Affe den Schalter losließ und zeigte einen unterschiedlichen Verlauf, abhängig davon, ob der Affe eine falsche oder eine richtige Entscheidung getroffen hatte.

Zusammenfassend zeigt diese Studie, daß Kategorien schon im sensorischen Cortex präsent sind. Die Assymetrie in die Repräsentation von drei untersuchten akustischen Karegorien und auch die Repräsentation von Reizen aus anderen Modalitäten sind vom Lernprozess initiiert und zeigen die hohe Lernplastizität vom auditorischen Cortex.

II. Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation mit dem Thema

„Firing activities of auditory cortical neurons during categorical task performance in behaving monkeys“

selbständig verfaßt, nicht schon als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr.rer.nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, den 30.05.2006

Diplom-Psychologin Elena Selezneva

III. List of publications

Brosch, M., Selezneva, E. and Scheich, H. (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25(29):6797-806.

Brosch, M., Selezneva, E., Bucks, C. and Scheich, H. (2004) Macaque monkeys discriminate pitch relationships. *Cognition.* 91(3):259-72.

Abstracts

Brosch, M., Oshurkova, E., Melikyan, A., Selezneva, E., Scheich, H. (2004) Reward-related firing in monkey auditory cortex. Abstract *Society for Neuroscience 34rd Annual Meeting,* San Diego. Program No. 486.10.

Selezneva E., Scheich H. and Brosch M. (2004) Context-dependent changes of the response properties of neurons in the auditory cortex of behaving monkeys. Abstract *4th Forum of European Neuroscience*, Lisbon. Program No. 226.13.

Brosch, M., Oshurkova, E., Melikyan, A., Selezneva, E. and Scheich, H. (2004) Contingency-related firing in the auditory cortex of monkeys performing an auditory discrimination task. Abstract *27th ARO Meeting*, Daytona Beach. No. 581, p. 195.

Brosch, M., Selezneva, E., Oshurkova, E., Melikyan, A., Goldschmidt, J. and Scheich, H. (2003) Neuronal activity in monkey auditory cortex during the performance of an auditory discrimination task. Abstract *Society for Neuroscience 33rd Annual Meeting,* New Orleans. Program No. 488.2.

Selezneva, E., Scheich H., Brosch M. (2003) Processing of sound sequences in the auditory cortex of behaving monkey. In *Proceedings of the International Conference on Auditory Cortex Magdeburg,* Shaker Verlag, Aachen, p.66.

Selezneva, E., Oshurkova, E., Scheich, H. and Brosch, M (2003) Comparison of the primary and the caudomedial field of monkey’s auditory cortex. In *Proceedings of the 29th Goettingen Neurobiology Conference and the 5th Meeting of German Neuroscience Society,* p.452-453.

Selezneva, E., Scheich, H., Budinger, E. and Brosch, M (2003) Spatial organisation of response properties in the primary and the caudomedial field of monkey's auditory cortex. Abstract *26th ARO Meeting*, Daytona Beach. No. 206, p. 53-54.

Brosch, M., Selezneva, E., Scheich, H. (2003) Discrimination of pitch direction by macaque monkeys. Abstract *26th ARO Meeting,* Daytona Beach. No. 295, p. 74-75.

IV. Curriculum vitae

Name: Elena Selezneva
Date of birth: July 05, 1977
Place of birth: Moscow, Russia
Sex: female
Citizenship: Russia

Education

2000	Master of science degree in Psychology from Moscow State University Diploma thesis: “Relationship between subjective dissimilarity and evoked potentials by perception of graphic faces”

Experience

9/1994 – 12/1994	Undergraduate student Faculty of Mechanics and Mathematics, Moscow State University, Moscow, Russia
9/1995 – 7/2000	Undergraduate student Department of Psychophysiology Faculty of Psychology, Moscow State University, Moscow, Russia
9/1999 –10/2000	Practical work in Laboratory of Neurophysiology Basis of Mind, Institute of Psychology of Russian Academy of Science, Moscow, Russia
from 10/2000	PhD Student Department Auditory Learning & Speech, Leibniz Institute for Neurobiology, Magdeburg, Germany